

R/Wiman, B.; Wallen, P.
 Eur. J. Biochem. 58, 539-547, 1975
 A/Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen the
 A/Reference number: A04626; MUID:76043692; PMID:126863
 A/Accession: A04626
 A/Molecule type: protein
 A/Residues: 483-507, E', 509-604 <W13>
 R/Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.
 J. Biol. Chem. 248, 1631-1633, 1973
 A/Title: The primary structure of human plasminogen. II. The histidine loop of human pla
 A/Reference number: A92125; MUID:73149248; PMID:4694729
 A/Contents: annotation; active site
 R/Groskopf, W.R.; Summaria, L.; Robbins, K.C.
 J. Biol. Chem. 244, 3590-3597, 1969
 A/Title: Studies on the active center of human plasmin. Partial amino acid sequence of a
 A/Reference number: A92048; MUID:69234739; PMID:4240117
 A/Contents: annotation; active site
 R/Trexler, M.; Vali, Z.; Pathy, L.
 J. Biol. Chem. 257, 7401-7406, 1982
 A/Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.
 A/Reference number: A92382; MUID:82213905; PMID:6919639
 A/Contents: annotation; omega-aminocarboxylic acid binding sites
 R/Vali, Z.; Pathy, L.
 J. Biol. Chem. 259, 13690-13694, 1984
 A/Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
 A/Reference number: A92458; MUID:85054794; PMID:6094526
 A/Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site
 R/Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marri, D.; Soehndel, S.; McCance, S.G.;
 J. Biol. Chem. 271, 29461-29467, 1996
 A/Title: Kringle domains of human angiotensin. Characterization of the anti-proliferativ
 A/Reference number: A58811; MUID:97067211; PMID:8910613
 A/Contents: annotation
 R/Lijnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.
 Biochemistry 37, 4699-4702, 1998
 A/Title: Generation of an angiotensin-like fragment from plasminogen by stromelysin-1 (M
 A/Reference number: A58812; MUID:9548733; PMID:9548733
 A/Contents: annotation
 R/Tulinsky, A.; Mulichak, A.M.
 submitted to the Brookhaven Protein Data Bank, July 1991
 A/Reference number: A51341; PDB:1PK4
 A/Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
 R/Tulinsky, A.; Wu, T.P.
 submitted to the Brookhaven Protein Data Bank, July 1991
 A/Reference number: A51488; PDB:2PK4
 A/Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454
 R/Wu, T.P.; Tulinsky, A.
 submitted to the Brookhaven Protein Data Bank, August 1993
 A/Reference number: A51911; PDB:1PKR
 A/Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
 R/Padmanabhan, K.; Tulinsky, A.
 submitted to the Brookhaven Protein Data Bank, April 1994
 A/Reference number: A52408; PDB:1PMK
 A/Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
 R/Tulinsky, A.; Mathews, I.I.
 submitted to the Brookhaven Protein Data Bank, December 1995
 A/Reference number: A65244; PDB:1CEA
 A/Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
 R/Tulinsky, A.; Mathews, I.I.
 submitted to the Brookhaven Protein Data Bank, December 1995
 A/Reference number: A65245; PDB:1CEB
 A/Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
 R/Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.
 Biochemistry 30, 10576-10588, 1991
 A/Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 A
 A/Reference number: A58819; MUID:92031502; PMID:1657148
 A/Contents: annotation
 R/Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.
 Biochemistry 30, 10589-10594, 1991
 A/Title: The refined structure of the epsilon-aminocaproic acid complex of human plasmin
 A/Reference number: A58818; MUID:92031503; PMID:1657149
 A/Contents: annotation
 R/de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, M.
 Biochemistry 31, 270-279, 1992

A/Title: Crystal structure of the kringle 2 domain of tissue plasminogen activ
 A/Reference number: A39483; MUID:92118803; PMID:1310033
 A/Contents: annotation; X-ray crystallography, 2.4 angstroms
 R/Stec, B.; Tester, M.M.; Whitlow, M.; Yamano, A.
 submitted to the Brookhaven Protein Data Bank, June 1995
 A/Reference number: A65980; PDB:1KRN
 A/Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-45
 R/Rejante, M.; Llinas, M.
 submitted to the Brookhaven Protein Data Bank, August 1996
 A/Reference number: A65803; PDB:1HPJ
 A/Contents: annotation; conformation by (1)H-NMR, residues 103-181
 R/Rejante, M.; Llinas, M.
 submitted to the Brookhaven Protein Data Bank, August 1996
 A/Reference number: A65804; PDB:1HPK
 A/Contents: annotation; conformation by (1)H-NMR, residues 103-181
 R/Rejante, M.R.; Llinas, M.
 Eur. J. Biochem. 221, 927-937, 1994
 A/Title: (1)H-NMR assignments and secondary structure of human plasminogen kri
 A/Reference number: 943645; MUID:94237157; PMID:8181475
 A/Contents: annotation; conformation by (1)H-NMR, residues 96-184
 R/Rejante, M.R.; Llinas, M.
 Eur. J. Biochem. 221, 939-949, 1994
 A/Title: Solution structure of the epsilon-aminohexanoic acid complex of human
 A/Reference number: A58817; MUID:94237158; PMID:8181476
 A/Contents: annotation; conformation by (1)H-NMR
 C/Comment: plasminogen is synthesized by the kidney and is present in plasma a
 d PIR:FGH383).
 C/Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:IFHU2) imme
 rg-580, resulting in two chains connected by two disulfide bonds. Without the
 C/Comment: Microplasmin is formed by autolytic cleavage of plasmin under artif
 C/Comment: Stromelysin 1 (see PIR:KCHUS1) acts on plasminogen to produce angio
 ting solid tumors.
 C/Genetics:
 A/Gene: GDB:PLG
 A/Cross-references: GDB:119498; OMIM:173350
 A/Map position: 6q26-6q27
 A/Introns: 17/1; 52/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2;
 C/Function:
 A/Description: dissolves the fibrin of blood clots; acts as a proteolytic fact
 ns the walls of the graafian follicle; also activates the urokinase-type plas
 A/Pathway: fibrinolysis
 C/Superfamily: plasmin; kringle homology; plasminogen-related protein precurs
 C/Keywords: angiotensin inhibitor; blood; duplication; fibrinolysis; glycopro
 F1-19/Domain: plasminogen-related protein precursor homology <PLPH>
 F1-19/Domain: signal sequence #status predicted <SIG>
 F20-810/Product: plasminogen #status experimental <PRO>
 F79-466/Product: angiotensin #status experimental <APT>
 F97-580,581-810/Product: plasmin #status experimental <MAT>
 F103-181/Domain: kringle homology <KR1>
 F185-262/Domain: kringle homology <KR2>
 F275-352/Domain: kringle homology <KR3>
 F377-454/Domain: kringle homology <KR4>
 F481-560/Domain: kringle homology <KR5>
 F550-580,581-810/Product: microplasmin #status experimental <MMT>

Query Match 22.2%; Score 501; DB 1; Length 810;
 Best Local Similarity 33.7%; Pred. No. 4.8e-31;
 Matches 137; Conservative 45; Mismatches 169; Indels 56; Gaps 15
 QY 10 NCDCLNGTCVSNKYFNIHWCNPKFKGGQ-----HCEIDKSKTCYEGNGH 56
 DB 428 NPDADKGPWCFTDPSVWEYCNLKCSGTEASVAPPPVLLPNVETPSEDCMFGNGK 487
 QY 57 FYRGKASTDTMGRCPLPWSATVLQTYHAHR----SDALQGLGKNCYNRPD-NRRRP 111
 DB 488 GYRGKRAITVTGTPQDWAA-----QEPHRSITPTETNPRAGLEK-NYCRNPDGVDGVP 541
 QY 112 WCYVVGKLPVQECNCHDCKLKCQCGKTKLRPR---FKIIGGFTTINQWPFAAI 168
 DB 542 WCYT-TNPRKLYDCVDPQCA--APSPDCGKQVSPKCKPGRVVGGCVAPHSPWQVSL 598

RESULT 23

B61545
plasmin (EC 3.4.21.7) precursor - sheep (fragments)
N/Alternate names: plasminogen
N/Contains: miniplasminogen
C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C/Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 17-Mar-1999
C/Accession: B61545; S28200
R/Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A/Title: Structural aspects of the plasminogen of various species.
A/Reference number: A61545; MUID:8905015; PMID:3168975
A/Accession: B61545
A/Molecule type: protein
A/Residues: 1-37;38-117 <SCH>
R/Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 5, 21-25, 1992
A/Title: Complete amino acid sequence of ovine miniplasminogen.
A/Reference number: S28200; MUID:93149995; PMID:1492092
A/Accession: S28200
A/Molecule type: protein
A/Residues: 118-460 <SC2>
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase; zymogen
F:1-37/38-117/118-460/Product: plasminogen (fragments) #status experimental <PRO>
F:1-37/Domain: activation peptide (fragment) #status experimental <AP>
F:38-117,118-230,231-460/Product: plasmin (fragments) #status experimental <MAT>
F:41-118/Domain: kringle homology <KR4>
F:118-460/Product: miniplasminogen #status experimental <MIN>
F:132-211/Domain: kringle homology <KR5>
F:226-460/Domain: plasmin chain B #status experimental <BCH>
F:231-453/Domain: trypsin homology <TRV>
F:272,315,410/Active site: His, Asp, Ser #status predicted

Query Match 22.2; Score 502; DB 2; Length 460;

Best Local Similarity 34.7; Pred. No. 2.2e-31;

Matches 131; Conservative 43; Mismatches 158; Indels 46; Gaps 14;

QY 28 IHW--CNCPKKFGGCHCISDKTCYEGNGHYRGKASTDMGRPCLPWNASATVLOQTVH 85
DQ 108 VRWFENLKKAPQAPSVENPEADCMGLGKGYRGKATTVAGVPCQENAA-----QEPH 162
QY 86 AH----RSDALQLGKHNKRNPD-NRRPWCYVQVLKPLVQECMVHPCADGKMKFC 140
DQ 163 RHGIFTPTNPRAGLEK-NYCRNPDGVNGPWCT-TNPKLFYCDIPQC---ESSFDC 217
QY 141 GQKTLRP--FKIIGFEFTIENQPFATYRRHGGSVTVYCGGLISPCWVISATHC 197
DQ 218 GPKVEPKKCPARVVGCVATPHSPWQVSRRSR-----EHFCGGTLISFEWLTAAHC 273
QY 198 FIDYPKKEDYTVYLGSRNLNSTQGMKFEVENILHKDYSADTFLAHNDIALKIRSK 257
DQ 274 LDSILGPSFTVILGAHYEMAREASVQEI PVSRFLPESRA-----DIALKLSP- 324
QY 258 GRCAQPSFTIOTICLPSMYNDPQF----GTSCETGTGKENSITDLYPEQLKMTVVKLIS 313
DQ 325 ---AVITDEVIPACLPSS-----PNYVADKTCYITGTGEGTGT--FGVGRLEKARLPVIE 375
QY 314 HRECOQPHYSGEYVTKMLCAADPQWKTDSCQSGGGLVCSLQGRMTLTGIVSWGRCGA 373
DQ 376 NKVCNRYEYLNKRVKSTELCAGDLAGGTDCQSGGGLVCFKDKVILQGVTSWGLGCA 435
QY 374 LKDRGVYTRVSHLPWI 391
DQ 436 RPNKPGVYRVSTVYPWI 453

RESULT 24

PLHU

plasmin (EC 3.4.21.7) precursor [validated] - human

N/Alternate names: plasminogen precursor [misnomer]

N/Contains: angiotatin; microplasma; plasminogen
C/Species: Homo sapiens (man)
C/Date: 24-Apr-1984 #sequence_revision 02-Dec-1994 #text_change 15-Sep-2000
C/Accession: A35229; I52242; A36646; I62738; I84609; S03735; A00929; A04627; A0
R/Petersen, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
J. Biol. Chem. 265, 6104-6111, 1990
A/Title: Characterization of the gene for human plasminogen, a key proenzyme in
A/Reference number: A35229; MUID:90202879; PMID:2318848
A/Accession: A35229
A/Molecule type: DNA
A/Residues: 1-810 <PET>
A/Cross-references: GB:J05286; GB:M34276; NID:G190064; PIDN:AAA60113.1; PID:G38
A/Experimental source: leukocyte; lung fibroblast
R/Margaret, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolengh
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
A/Title: Definition of the transcription initiation site of human plasminogen S
A/Reference number: I52242; MUID:91097523; PMID:2268308
A/Accession: I52242
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-16 <MAL1>
A/Cross-references: GB:M62890; NID:G190092; PIDN:AAA36454.1; PID:G553613
R/Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FEBS Lett. 213, 254-260, 1987
A/Title: Molecular cloning and characterization of a full-length cDNA clone for
A/Reference number: A26646; MUID:87162490; PMID:3030813
A/Accession: A26646
A/Molecule type: mRNA
A/Residues: 1-471, 'D', 473-810 <FOR>
A/Cross-references: GB:X05199; NID:G35530; PIDN:CAA28931.1; PID:G35531
A/Experimental source: liver
R/Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A/Title: Characterization of a complementary deoxyribonucleic acid coding for h
A/Reference number: I45961; MUID:85023311; PMID:6148961
A/Accession: I62738
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 292-471, 'D', 473-810 <MAL2>
A/Cross-references: GB:K02922; NID:G190112; PIDN:AAA60124.1; PID:G387031
A/Accession: I84609
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 367-419 <MAL3>
A/Cross-references: GB:K02921; NID:G190110; PIDN:AAA60123.1; PID:G190111
R/Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manne
Eur. J. Biochem. 114, 465-470, 1981
A/Title: Comparison of the primary structure of the N-terminal CNBr fragments o
A/Reference number: S03735; MUID:81212097; PMID:7238497
A/Accession: S03735
A/Molecule type: protein
A/Residues: 20-71, 'E', 73-76 <BRU>
R/Settrup-Jensen, L.; Petersen, T.E.; Magnusson, S.
submitted to the Atlas, July 1977
A/Reference number: A00929
A/Accession: A00929
A/Molecule type: protein
A/Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>
R/Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A/Title: Primary structure of the B-chain of human plasmin.
A/Reference number: A04627; MUID:77225245; PMID:142009
A/Accession: A04627
A/Molecule type: protein
A/Residues: 581-810 <WIL>
R/Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A/Title: Structural relationship between "glutamic acid" and "lysine" forms of
A/Reference number: A04625; MUID:75093329; PMID:122932
A/Accession: A04625
A/Molecule type: protein
A/Residues: 20-50, 'Q', 51-71, 'E', 73-85, 87-100 <WIL2>

N:Alternate names: Hageman factor (activated)
C/Species: Bos primigenius taurus (cattle)
C/Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C/Accession: S45281; A61329
R/Shibuya, Y.; Senba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A/Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): comparison of the heavy chain of porcine plasmin. Comparison of the heavy chain of porcine plasmin. Comparison of the heavy chain of porcine plasmin.
A/Reference number: S45281; MUID:94242782; PMID:8186251
A/Accession: S45281
A/Molecule type: mkNA
A/Residues: 1-593 <SHR>
A/Cross-references: GB:S70164
A/Note: The authors translated the codon GAG for residue 23 as Val, GAG for residue 70 as Pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 as Glu, and ATC for residue 505 as Leu
R/Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A/Title: Isolation and characterization of bovine factor XII (Hageman factor).
A/Reference number: A61329; MUID:77182112; PMID:861210
A/Accession: A61329
A/Molecule type: protein
A/Residues: 10-16, 'X', 18-19, 525-550 <FUJ>
C/Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology; C/Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma; F/37-78/Domain: fibronectin type II repeat homology <IF2>
F/88-120/Domain: EGF homology <EGF>
F/125-160/Domain: fibronectin type I repeat homology <FBI>
F/207-287/Domain: kringle homology <KR>
F/350-587/Domain: trypsin homology <TRY>
F/541/Active site: Ser #status predicted

Query Match 28.3%; Score 638; DB 2; Length 593;
Best Local Similarity 34.2%; Pred. No. 7.9e-42;
Matches 151; Conservative 59; Mismatches 164; Indels 68; Gaps 13;
6 QVPSNCPLNGGTCVSNKYFSNTHWNCNPKFGQHCIDKSKTCYE--GNHGFYRGKAS 63
166 QVCTNPLNGDSCLOAE---GHRLCRCAPSFAGRLCDVLDKASCYDDRRDGLSYRGMA 222
64 TDWGRCLPWN SATVLQOTY-HAHRSDALQLGLGKINYCRNPNDRRRPCVYQVGLKPL 122
223 TTLSGAPCQSWAS---EATVNVVTAQVLNWLGDHAFRCNPDNDTRPCWFFWKGDRLS 278
123 VQECWHDG--ADGKLKQ-----CGQ 142
279 WNYCLAPCAQAAGHEHFLPSPSALQKPESTTQTPLPSTSGWCSTPLASGPGGCGQ 338
143 ---KTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGSLISPCWVISATHCFI 199
339 RLKWLSSLNRVVGVLPAHPYIALYDQ-----HFCAGSLIAPCNVLTAAHCLQ 392
200 DYPKEDYIYVLRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIR-SKEG 258
393 NRPAKELTVVLDGDRHNSCEOCQTLAVRDYRLHEAFSPITYQH--DLALVRLQESADG 450
259 RCAQPSRTIOTICLPNMYNDPQFQTS--CEITGFGKNSDLYPEQLKMTVVVKLISHRE 316
451 CCAHPSFVQVCLPSTAAPASEAAVCEVAGHGFEGE--YSSFLQAQVPLIDPQR 509
317 CQPHYGVSEVTTKMLCAADPQWKTDCQSDSGGLVLC---SLQGRMTLGIYSWGRGCA 373
510 CSAPDVHGAFTQMLCAGLEGGTACQSDSGGLVLCEDETPERQILRGIYSWGS GCG 569
374 LKDKPGVTVRSHPLPIRSH 395
570 NLRKPGVTVDVANLWIREHT 591

RESULT 22
PLPG
plasmin (EC 3.4.21.7) precursor - pig (fragment)
N:Alternate names: plasminogen
N/Contains: miniplasminogen

C/Species: Sus scrofa domestica (domestic pig)
C/Date: 07-Sep-1990 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
C/Accession: S03733; A25834
R/Schaller, J.; Marti, T.; Roesselet, S.J.; Kaempfer, U.; Rickli, E.E.
Fibrinolysis 1, 91-102, 1987
A/Title: Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the heavy chain of porcine plasmin. Comparison of the heavy chain of porcine plasmin.
A/Reference number: S03733
A/Accession: S03733
A/Molecule type: protein
A/Residues: 1-560 <SCH>
R/Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manuerg, J.
Eur. J. Biochem. 114, 465-470, 1981
A/Title: Comparison of the primary structure of the N-terminal CNBr fragments of plasminogen and plasmin.
A/Reference number: S03735; MUID:81212097; PMID:7238497
A/Accession: S03737
A/Molecule type: protein
A/Residues: 1-57 <BRU>
R/Marti, T.; Schaller, J.; Rickli, E.E.
Eur. J. Biochem. 149, 279-285, 1985
A/Title: Determination of the complete amino-acid sequence of porcine miniplasminogen.
A/Reference number: A25834; MUID:85203907; PMID:3846533
A/Accession: A25834
A/Molecule type: protein
A/Residues: 450-790 <MAR>
C/Function:
A/Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in the walls of the graafian follicle; also activates the urokinase-type plasminogen pathway; fibrinolysis
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor; C/Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma; serine protease
F/1-790/Product: plasminogen #status predicted <PRO>
F/1-77/Domain: plasminogen-related protein precursor homology (fragment) <PLP>
F/78-560/Product: activation peptide #status predicted <APT>
F/84-162/Domain: kringle homology <KR1>
F/166-243/Domain: kringle homology <KR2>
F/256-333/Domain: kringle homology <KR3>
F/358-435/Domain: kringle homology <KR4>
F/450-790/Product: miniplasminogen #status experimental <MIN>
F/461-540/Domain: kringle homology <KR5>
F/561-790/Product: plasmin chain B #status experimental <BCH>
F/561-783/Domain: trypsin homology <TRY>
F/30-54,34-42,84-162,105-145,133-157,166-243,169-297,187-226,215-238,256-333,338-416,417-418,419-420,421-422,423-424,425-426,427-428,429-430,431-432,433-434,435-436,437-438,439-440,441-442,443-444,445-446,447-448,449-450,451-452,453-454,455-456,457-458,459-460,461-462,463-464,465-466,467-468,469-470,471-472,473-474,475-476,477-478,479-480,481-482,483-484,485-486,487-488,489-490,491-492,493-494,495-496,497-498,499-500,501-502,503-504,505-506,507-508,509-510,511-512,513-514,515-516,517-518,519-520,521-522,523-524,525-526,527-528,529-530,531-532,533-534,535-536,537-538,539-540,541-542,543-544,545-546,547-548,549-550,551-552,553-554,555-556,557-558,559-560,561-562,563-564,565-566,567-568,569-570,571-572,573-574,575-576,577-578,579-580,581-582,583-584,585-586,587-588,589-590,591-592,593-594,595-596,597-598,599-600,601-602,603-604,605-606,607-608,609-610,611-612,613-614,615-616,617-618,619-620,621-622,623-624,625-626,627-628,629-630,631-632,633-634,635-636,637-638,639-640,641-642,643-644,645-646,647-648,649-650,651-652,653-654,655-656,657-658,659-660,661-662,663-664,665-666,667-668,669-670,671-672,673-674,675-676,677-678,679-680,681-682,683-684,685-686,687-688,689-690,691-692,693-694,695-696,697-698,699-700,701-702,703-704,705-706,707-708,709-710,711-712,713-714,715-716,717-718,719-720,721-722,723-724,725-726,727-728,729-730,731-732,733-734,735-736,737-738,739-740,741-742,743-744,745-746,747-748,749-750,751-752,753-754,755-756,757-758,759-760,761-762,763-764,765-766,767-768,769-770,771-772,773-774,775-776,777-778,779-780,781-782,783-784,785-786,787-788,789-790,791-792,793-794,795-796,797-798,799-800,801-802,803-804,805-806,807-808,809-810,811-812,813-814,815-816,817-818,819-820,821-822,823-824,825-826,827-828,829-830,831-832,833-834,835-836,837-838,839-840,841-842,843-844,845-846,847-848,849-850,851-852,853-854,855-856,857-858,859-860,861-862,863-864,865-866,867-868,869-870,871-872,873-874,875-876,877-878,879-880,881-882,883-884,885-886,887-888,889-890,891-892,893-894,895-896,897-898,899-900,901-902,903-904,905-906,907-908,909-910,911-912,913-914,915-916,917-918,919-920,921-922,923-924,925-926,927-928,929-930,931-932,933-934,935-936,937-938,939-940,941-942,943-944,945-946,947-948,949-950,951-952,953-954,955-956,957-958,959-960,961-962,963-964,965-966,967-968,969-970,971-972,973-974,975-976,977-978,979-980,981-982,983-984,985-986,987-988,989-990,991-992,993-994,995-996,997-998,999-1000,1001-1002,1003-1004,1005-1006,1007-1008,1009-1010,1011-1012,1013-1014,1015-1016,1017-1018,1019-1020,1021-1022,1023-1024,1025-1026,1027-1028,1029-1030,1031-1032,1033-1034,1035-1036,1037-1038,1039-1040,1041-1042,1043-1044,1045-1046,1047-1048,1049-1050,1051-1052,1053-1054,1055-1056,1057-1058,1059-1060,1061-1062,1063-1064,1065-1066,1067-1068,1069-1070,1071-1072,1073-1074,1075-1076,1077-1078,1079-1080,1081-1082,1083-1084,1085-1086,1087-1088,1089-1090,1091-1092,1093-1094,1095-1096,1097-1098,1099-1100,1101-1102,1103-1104,1105-1106,1107-1108,1109-1110,1111-1112,1113-1114,1115-1116,1117-1118,1119-1120,1121-1122,1123-1124,1125-1126,1127-1128,1129-1130,1131-1132,1133-1134,1135-1136,1137-1138,1139-1140,1141-1142,1143-1144,1145-1146,1147-1148,1149-1150,1151-1152,1153-1154,1155-1156,1157-1158,1159-1160,1161-1162,1163-1164,1165-1166,1167-1168,1169-1170,1171-1172,1173-1174,1175-1176,1177-1178,1179-1180,1181-1182,1183-1184,1185-1186,1187-1188,1189-1190,1191-1192,1193-1194,1195-1196,1197-1198,1199-1200,1201-1202,1203-1204,1205-1206,1207-1208,1209-1210,1211-1212,1213-1214,1215-1216,1217-1218,1219-1220,1221-1222,1223-1224,1225-1226,1227-1228,1229-1230,1231-1232,1233-1234,1235-1236,1237-1238,1239-1240,1241-1242,1243-1244,1245-1246,1247-1248,1249-1250,1251-1252,1253-1254,1255-1256,1257-1258,1259-1260,1261-1262,1263-1264,1265-1266,1267-1268,1269-1270,1271-1272,1273-1274,1275-1276,1277-1278,1279-1280,1281-1282,1283-1284,1285-1286,1287-1288,1289-1290,1291-1292,1293-1294,1295-1296,1297-1298,1299-1300,1301-1302,1303-1304,1305-1306,1307-1308,1309-1310,1311-1312,1313-1314,1315-1316,1317-1318,1319-1320,1321-1322,1323-1324,1325-1326,1327-1328,1329-1330,1331-1332,1333-1334,1335-1336,1337-1338,1339-1340,1341-1342,1343-1344,1345-1346,1347-1348,1349-1350,1351-1352,1353-1354,1355-1356,1357-1358,1359-1360,1361-1362,1363-1364,1365-1366,1367-1368,1369-1370,1371-1372,1373-1374,1375-1376,1377-1378,1379-1380,1381-1382,1383-1384,1385-1386,1387-1388,1389-1390,1391-1392,1393-1394,1395-1396,1397-1398,1399-1400,1401-1402,1403-1404,1405-1406,1407-1408,1409-1410,1411-1412,1413-1414,1415-1416,1417-1418,1419-1420,1421-1422,1423-1424,1425-1426,1427-1428,1429-1430,1431-1432,1433-1434,1435-1436,1437-1438,1439-1440,1441-1442,1443-1444,1445-1446,1447-1448,1449-1450,1451-1452,1453-1454,1455-1456,1457-1458,1459-1460,1461-1462,1463-1464,1465-1466,1467-1468,1469-1470,1471-1472,1473-1474,1475-1476,1477-1478,1479-1480,1481-1482,1483-1484,1485-1486,1487-1488,1489-1490,1491-1492,1493-1494,1495-1496,1497-1498,1499-1500,1501-1502,1503-1504,1505-1506,1507-1508,1509-1510,1511-1512,1513-1514,1515-1516,1517-1518,1519-1520,1521-1522,1523-1524,1525-1526,1527-1528,1529-1530,1531-1532,1533-1534,1535-1536,1537-1538,1539-1540,1541-1542,1543-1544,1545-1546,1547-1548,1549-1550,1551-1552,1553-1554,1555-1556,1557-1558,1559-1560,1561-1562,1563-1564,1565-1566,1567-1568,1569-1570,1571-1572,1573-1574,1575-1576,1577-1578,1579-1580,1581-1582,1583-1584,1585-1586,1587-1588,1589-1590,1591-1592,1593-1594,1595-1596,1597-1598,1599-1600,1601-1602,1603-1604,1605-1606,1607-1608,1609-1610,1611-1612,1613-1614,1615-1616,1617-1618,1619-1620,1621-1622,1623-1624,1625-1626,1627-1628,1629-1630,1631-1632,1633-1634,1635-1636,1637-1638,1639-1640,1641-1642,1643-1644,1645-1646,1647-1648,1649-1650,1651-1652,1653-1654,1655-1656,1657-1658,1659-1660,1661-1662,1663-1664,1665-1666,1667-1668,1669-1670,1671-1672,1673-1674,1675-1676,1677-1678,1679-1680,1681-1682,1683-1684,1685-1686,1687-1688,1689-1690,1691-1692,1693-1694,1695-1696,1697-1698,1699-1700,1701-1702,1703-1704,1705-1706,1707-1708,1709-1710,1711-1712,1713-1714,1715-1716,1717-1718,1719-1720,1721-1722,1723-1724,1725-1726,1727-1728,1729-1730,1731-1732,1733-1734,1735-1736,1737-1738,1739-1740,1741-1742,1743-1744,1745-1746,1747-1748,1749-1750,1751-1752,1753-1754,1755-1756,1757-1758,1759-1760,1761-1762,1763-1764,1765-1766,1767-1768,1769-1770,1771-1772,1773-1774,1775-1776,1777-1778,1779-1780,1781-1782,1783-1784,1785-1786,1787-1788,1789-1790,1791-1792,1793-1794,1795-1796,1797-1798,1799-1800,1801-1802,1803-1804,1805-1806,1807-1808,1809-1810,1811-1812,1813-1814,1815-1816,1817-1818,1819-1820,1821-1822,1823-1824,1825-1826,1827-1828,1829-1830,1831-1832,1833-1834,1835-1836,1837-1838,1839-1840,1841-1842,1843-1844,1845-1846,1847-1848,1849-1850,1851-1852,1853-1854,1855-1856,1857-1858,1859-1860,1861-1862,1863-1864,1865-1866,1867-1868,1869-1870,1871-1872,1873-1874,1875-1876,1877-1878,1879-1880,1881-1882,1883-1884,1885-1886,1887-1888,1889-1890,1891-1892,1893-1894,1895-1896,1897-1898,1899-1900,1901-1902,1903-1904,1905-1906,1907-1908,1909-1910,1911-1912,1913-1914,1915-1916,1917-1918,1919-1920,1921-1922,1923-1924,1925-1926,1927-1928,1929-1930,1931-1932,1933-1934,1935-1936,1937-1938,1939-1940,1941-1942,1943-1944,1945-1946,1947-1948,1949-1950,1951-1952,1953-1954,1955-1956,1957-1958,1959-1960,1961-1962,1963-1964,1965-1966,1967-1968,1969-1970,1971-1972,1973-1974,1975-1976,1977-1978,1979-1980,1981-1982,1983-1984,1985-1986,1987-1988,1989-1990,1991-1992,1993-1994,1995-1996,1997-1998,1999-2000,2001-2002,2003-2004,2005-2006,2007-2008,2009-2010,2011-2012,2013-2014,2015-2016,2017-2018,2019-2020,2021-2022,2023-2024,2025-2026,2027-2028,2029-2030,2031-2032,2033-2034,2035-2036,2037-2038,2039-2040,2041-2042,2043-2044,2045-2046,2047-2048,2049-2050,2051-2052,2053-2054,2055-2056,2057-2058,2059-2060,2061-2062,2063-2064,2065-2066,2067-2068,2069-2070,2071-2072,2073-2074,2075-2076,2077-2078,2079-2080,2081-2082,2083-2084,2085-2086,2087-2088,2089-2090,2091-2092,2093-2094,2095-2096,2097-2098,2099-2100,2101-2102,2103-2104,2105-2106,2107-2108,2109-2110,2111-2112,2113-2114,2115-2116,2117-2118,2119-2120,2121-2122,2123-2124,2125-2126,2127-2128,2129-2130,2131-2132,2133-2134,2135-2136,2137-2138,2139-2140,2141-2142,2143-2144,2145-2146,2147-2148,2149-2150,2151-2152,2153-2154,2155-2156,2157-2158,2159-2160,2161-2162,2163-2164,2165-2166,2167-2168,2169-2170,2171-2172,2173-2174,2175-2176,2177-2178,2179-2180,2181-2182,2183-2184,2185-2186,2187-2188,2189-2190,2191-2192,2193-2194,2195-2196,2197-2198,2199-2200,2201-2202,2203-2204,2205-2206,2207-2208,2209-2210,2211-2212,2213-2214,2215-2216,2217-2218,2219-2220,2221-2222,2223-2224,2225-2226,2227-2228,2229-2230,2231-2232,2233-2234,2235-2236,2237-2238,2239-2240,2241-2242,2243-2244,2245-2246,2247-2248,2249-2250,2251-2252,2253-2254,2255-2256,2257-2258,2259-2260,2261-2262,2263-2264,2265-2266,2267-2268,2269-2270,2271-2272,2273-2274,2275-2276,2277-2278,2279-2280,2281-2282,2283-2284,2285-2286,2287-2288,2289-2290,2291-2292,2293-2294,2295-2296,2297-2298,2299-2300,2301-2302,2303-2304,2305-2306,2307-2308,2309-2310,2311-2312,2313-2314,2315-2316,2317-2318,2319-2320,2321-2322,2323-2324,2325-2326,2327-2328,2329-2330,2331-2332,2333-2334,2335-2336,2337-2338,2339-2340,2341-2342,2343-2344,2345-2346,2347-2348,2349-2350,2351-2352,2353-2354,2355-2356,2357-2358,2359-2360,2361-2362,2363-2364,2365-2366,2367-2368,2369-2370,2371-2372,2373-2374,2375-2376,2377-2378,2379-2380,2381-2382,2383-2384,2385-2386,2387-2388,2389-2390,2391-2392,2393-2394,2395-2396,2397-2398,2399-2400,2401-2402,2403-2404,2405-2406,2407-2408,2409-2410,2411-2412,2413-2414,2415-2416,2417-2418,2419-2420,2421-2422,2423-2424,2425-2426,2427-2428,2429-2430,2431-2432,2433-2434,2435-2436,2437-2438,2439-2440,2441-2442,2443-2444,2445-2446,2447-2448,2449-2450,2451-2452,2453-2454,2455-2456,2457-2458,2459-2460,2461-2462,2463-2464,2465-2466,246

```

Db 410 AAHCLQDRPAPEDTVVLGQRRNHSCEPOTLAVRSYRLHEAFS--PVSQYQHDALLRL 467
Qy 254 R-SKEGCAQPSRTIQTICLPSMYNDPQFGTSCBITGKXENSDYLYPQLKMTVVVKLI 312
Db 468 QEDADGSCALLSPVQVCLPSGAARPSSETTLQCVAGWGHCQFEGAEYASFLQEAQVPL 527
Qy 313 SHRECOQPHYGVSEVTTKMLCAADPQWKTSCQDGGSLPVCSLQ---RMTLTGIVSWG 369
Db 528 SLERCSAPDVHGSSILPQMLCAGLEGGTDACQDGGSLPVCEQDAERELTLQGIISWG 587
Qy 370 RGCALKDKPGVTVVSHPLPWIRGHT 395
Db 588 SGCGDRNPGVYTDVAVYLAMIREHT 613

RESULT 19
JC5878
plasma hyaluronan-binding protein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999
C:Accession: JC5878
R:Hashimoto, K.; Tobe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, N.
  Biol. Pharm. Bull. 20, 1127-1130, 1997
A:Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-binding
  protein precursor
A:Reference number: JC5878; MUID:98065239; PMID:9401717
A:Accession: JC5878
A:Molecule type: mRNA
A:Residues: 1-558 <HAS>
C:Comment: This protein acts as serine protease.
C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin
  F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <MATH>
F:75-106/Domain: EGF homology <EG1>
F:113-145/Domain: EGF homology <EG2>
F:152-185/Domain: EGF homology <EG3>
F:192-274/Domain: kringle homology <KRI>
F:312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MATH>
F:312-548/Domain: trypsin homology <TRY>

Query Match 29.4%; Score 664.5; DB 2; Length 558;
Best Local Similarity 36.9%; Pred. No. 6.5e-44;
Matches 154; Conservative 57; Mismatches 163; Indels 43; Gaps 14;

Qy 13 CLNGGTCVSNKYFNSNIHWCNCPKFGQHCIDSKTCYEGNGHFYRGKASTDTMGRPCL 72
Db 157 CQNGVCSSRRRRSRF--TCACPDQYKGFCEIGPD--DCYVGDGYSYRGVSKTYNQNPCL 214
Qy 73 PWSATVLOQTVAHRS DALQLGLGKNYCRNPNRRPWCYVOVGLKPL-----VQBC 126
Db 215 YWNSHLLQETYNMFEDAE THGIAEHNF CNPDGDH KPCFVKVNSEKVKWEYCDVTVC 274
Qy 127 MYHDCADGKLKPF-----CQGQKTLRPR--FKIIGGEFTTIENQPFPAIY---- 169
Db 275 PVPDTPNVPESLLEPVMELPQFESCGKTEVAHAKVIRYGGFKSTAGKHPQVSLQSLP 334
Qy 170 ---RRHSGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKF 226
Db 335 LTTSMPQG---HFCCGALIHPCWVLTAAHC--TDINTKLVV-LGDDQLKKTSHQTF 389
Qy 227 EVENILHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDP--QFGTSC 285
Db 389 RVEKILKYQXNERDEIPHNDIALKLPVGGHCALESRYKVTCLPS---DPPPSGTEC 445
Qy 286 EITGFGKENS TDLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTKMLCAADPQWKT--TDSC 344
Db 446 HISGGMVETG--GSRQLLDKAKVLIANPLCNLSRLQYDHTIDDSMI CAGNLQKPGSDTC 503
Qy 345 QGDSGGLPVCSLQGRMTLTGIVSWGRCALKDKPGVTVVSHPLPWIRSHTKKEENG 401
Db 504 QGDSGGLPFCXKDGTVYVYGVSWGQECG--KKPGVYTVQVTKFLNWKTIKESG 558

```

```

RESULT 20
JC4795
plasma hyaluronan-binding protein precursor - human
N:Alternate names: hepatocyte growth factor activator-like protein; PHBP
N:Contains: serine proteinase (EC 3.4.21.-)
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 19-Jul-2002
C:Accession: JC4795
R:Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, T.; Tomi, M.
  J. Biochem. 119, 1157-1165, 1996
A:Title: Purification and characterization of a novel hyaluronan-binding protein
  precursor
A:Reference number: JC4795; MUID:96425001; PMID:8827452
A:Accession: JC4795
A:Molecule type: mRNA
A:Residues: 1-560 <CHO>
A:Cross-references: GB:S83182; MID:gl836158; PIDN:AAB45909.1; PID:gl836159
A:Experimental source: plasma
A:Note: parts of this sequence, including the amino ends of the mature chains,
  are determined from the same precursor
C:Genetics:
A:Gene: GDB:HABP2; HABP; PHBP; HGFA1
A:Cross-references: GDB:4573562
C:Complex: a disulfide-bonded heterodimer of chains produced from the same precursor
C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology
C:Keywords: chondroitin sulfate proteoglycan; glycoprotein; hyaluronic acid; hy
  F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-313/Product: plasma hyaluronan-binding protein, 50K chain #status predicted
  F:77-108/Domain: EGF homology <EG1>
F:115-147/Domain: EGF homology <EG2>
F:154-187/Domain: EGF homology <EG3>
F:194-276/Domain: trypsin homology <KRI>
F:314-550/Domain: kringle homology <KRI>
F:314-516/Product: plasma hyaluronan-binding protein, catalytic chain #status p
  F:54-207/Binding site: carboxylate (Asn) (covalent) #status predicted
F:77-88,82-97,99-108,115-125,138-147,154-165,159-176,178-187,194-276,21
  F:362,405,509/Active site: His, Asp, Ser #status predicted

Query Match 28.9%; Score 651.5; DB 1; Length 560;
Best Local Similarity 36.4%; Pred. No. 6.6e-43;
Matches 152; Conservative 58; Mismatches 161; Indels 47; Gaps 14;

Qy 13 CLNGGTCVSNKYFNSNIHWCNCPKFGQHCIDSKTCYEGNGHFYRGKASTDTMGRPCL 72
Db 159 CQNGATCSRHRKRSKF--TCACPDQKGFCEIG--SDCCVYGDGYSYRGGMNRTVNOHACL 216
Qy 73 PWSATVLOQTVAHRS DALQLGLGKNYCRNPNRRPWCYVOVGLKPLVQE--CMWHDC 131
Db 217 YWNSHLLQETYNMFEDAE THGIEHNF CNPDADKPCWCFIKVTNDKVKWEYCDVSAC 276
Qy 132 ADGKLKPF-----CQGQKTLRPR--FKIIGGEFTTIENQPFPAIYRRHR 173
Db 277 SAQDVAYPESPTSTKLPFGDSCGKTEIAERKIKRIYGGFKSTAGKHPWQASLQ---- 332
Qy 174 GGSVT-----YVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMK 225
Db 333 -SSLPLTISMPQGHFCGALIHPCWVLTAAHC--TDI--KTRHLKVLGDDQLKKEEFHEQS 389
Qy 226 FEVENILHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPQF--QT 283
Db 390 FRVEKIFKSHYNERDEIPHNDIALKLPVGGHCALESRYKVTCLP---DGSPPSSG 445
Qy 284 SCEITGFGKENS TDLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTKMLCAADPQWKT--TD 342
Db 446 ECHISGMVETGK--GSRQLLDKAKVLIANPLCNLSRLQYDHTIDDSMI CAGNLQKPGQD 503
Qy 343 SCQDGGGLPVCSLQGRMTLTGIVSWGRCALKDKPGVTVVSHPLPWIRSHTKKEENG 400
Db 504 TCQDGGGLPFCXKDGTVYVYGVSWGLECG--KKPGVYTVQVTKFLNWKTIKESG 559

RESULT 21
S45281
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)

```

Query Match 31.0%; Score 700.5; DB 2.; Length 603;
Best Local Similarity 36.7%; Pred No. 1.le-46;
Matches 160; Conservative 60; Mismatches 147; Indels 69; Gaps 13;

Qy 13 CLNGGTCVSNKYFNIHWCNCKFKGGOHCEIDKSKTCYEGNHFFYRGKASTDTMGPRCL 72
||| |
Db 182 CLNGGRCL---VEGHLLCDPCMGYTGPFDLDTASCYEGRGVSYRGMARTTVSGAKCQ 238
||| |
Qy 73 PWSNATVLQOITYAHRS--ALQLGLGKHNYCRNPDRRRPWCYYOVVGLKPLVOECMVHDC 131
||| |
Db 239 RWAS-----EATYRNNTAEQALRGHLGHTFCENPDNDTRPMCFVMWGNRLSWEYCDLAQC 294
||| |
Qy 132 -----ADGKLKFQ-----CGQKTLPFRF 149
||| |
Db 295 QYPQPATPHDRPEHPKLPSSRLSILOTPQTTONQALANELPETSSLCCGOR-LRKRL 353
||| |
Qy 150 ---KIIGGEFTTIENPFAAIYRRHRGGSVTVCGSLISPCWISATHCFDVPKE 205
||| |
Db 354 SSLSRIVGLVALPCAHPYAALY----WGS-NFGSGSLIAPCWLTAAHCLONREPAE 407
||| |
Qy 206 DYIVLGRSLNSTQEMKEFEVENLIHKDYSADTLAHDNDIALKI-RSKGRCAQPS 264
||| |
Db 408 ELKVYLGDQRNQSCHECHQTLAVHSYKLHEAFSP--SYLNDLALLRLQSADGSCAQLS 465
||| |
Qy 265 RTIOTICLPSMYNDPPQG--TSCITGFGENSTDYLYPEOLKMTVVVKLSHRECQQPHY 322
||| |
Db 466 PYVQTVCLPSPAPPSESTTCCEVAGWGHCFCGABEYSFLOEAQVPLISSRCSSPEV 525
||| |
Qy 323 YGSEVTTYKLCADPQWKTSQGDSSGGPLV---SLQGRMTLTGIVSWGRCALKDKPKG 379
||| |
Db 526 HGDAFLGMLCAGFLGEGTDACQDSGGPLVCEDEAAEHRLILRGIVSWGSGCDNRNKP 585
||| |
Qy 380 VYTRVSFLPWIRSH 395
||| |
Db 586 VYTDVASVLTWIKHT 601
||| |

RESULT 18
KfHu12
coagulation factor Xira (EC 3.4.21.38) precursor [validated] - human
N/Alternate names: Hageman factor (activated)
C/Species: Homo sapiens (man)
C/Date: 27-Nov-1985 #sequence revision 30-Jun-1991 #text change 08-Dec-2000
C/Accession: A29411; A26814; J00930; A25191; A22248; A21037
P/Cool, D.E.; Macgillivray, R.T.A.
J. Biol. Chem. 262, 13662-13673, 1987
A/Title: Characterization of the human blood coagulation factor XII (Hageman).
A/Reference number: A29411; MUID:88007593; PMID:2888762
A/Accession: A29411
A/Molecule type: DNA
A/Residues: 1-615 <COO>
A/Cross-references: GB:M17466; GB:J02807; NID:g180355; PIDN:AAB59490.1; PID:g180357
P/Tripodi, M.; Citarista, F.; Guida, S.; Galeffi, P.; Fantoni, A.; Cortese, R.
Nucleic Acids Res. 14, 3146, 1986
A/Title: cDNA sequence coding for human coagulation factor XII (Hageman).
A/Reference number: A26814; MUID:96176794; PMID:3754331
A/Accession: A26814
A/Molecule type: mRNA
A/Residues: 4-615 <TRI>
A/Cross-references: GB:X13135; NID:g182231; PIDN:AAA70225.1; PID:g182292
P/Cool, D.E.; Edgell, C.J.S.; Louie, G.V.; Zoller, M.J.; Brayer, G.D.; Macgillivray, R.T.
J. Biol. Chem. 260, 13666-13676, 1985
A/Title: Characterization of human blood coagulation factor XII cdna. Prediction of the
A/Reference number: A00930; MUID:96033830; PMID:3877053
A/Accession: A00930
A/Molecule type: mRNA
A/Residues: 14-332,'S',334-615 <CO2>
A/Cross-references: GB:M1723; NID:g180358; PIDN:AAA51986.1; PID:g180359
P/Que, B.G.; Davis, E.W.
Biochemistry 25, 1525-1528, 1986
A/Title: Characterization of a cDNA coding for human factor XII (Hageman factor).
A/Reference number: A25191; MUID:96216049; PMID:3011063

A:Accession: A25191
A:Molecule type: mRNA
A:Residues: 146-378; G*, 380-615 <QUE>
A:Cross-references: GB:M13147; NID:G180360; PIDN:AAA70224.1; PFD:G180361
R:McMullen, B.A.; Fujikawa, K.
J. Biol. Chem. 260, 5328-5341, 1985
A>Title: Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated)
A:Reference number: A22248; MUID:85182674; PMID:3886654
A:Accession: A22248
A:Molecule type: protein
A:Residues: 20-379 <NCMC>
R:Fujikawa, K.; McMullen, B.A.
J. Biol. Chem. 258, 10924-10933, 1983
A>Title: Amino acid sequence of human beta-factor XIla.
A:Reference number: A21037; MUID:83291041; PMID:6604055
A:Accession: A21037
A:Molecule type: protein
A:Residues: 354-362;373-615 <FUJ>
R:Harris, R.J.; Ling, V.T.; Spellman, M.W.
J. Biol. Chem. 267, 5102-5107, 1992
A>Title: O-linked fucose is present in the first epidermal growth factor domain of factor XIIIa.
A:Reference number: A44606; MUID:92184750; PMID:1544894
A:Contents: annotation; carbohydrate binding site
C:Genetics:
A:Gene: GDB:F12
A:Cross-references: GDB:I19892; OMIM:234000
A:Map position: 5q34-5qter
A:Introns: 19/3; 39/1; 72/2; 96/1; 133/1; 177/1; 212/1; 267/2; 340/1; 417/2; 471/5
C:Complex: factor XII, prekallikrein, and HMW kininogen form a complex bound to fibrin monomer.
C:Function:
A:Description: factor XIIa catalyzes the proteolytic activation of plasminogen to plasmin.
A:Pathway: blood coagulation; fibrinolysis
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat
C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; kringle; plasma
F:1-19/Domain: signal sequence #status predicted <Sig>
F:20-372,373-615/Product: coagulation factor XIIa, alpha form #status experimentally confirmed
F:47-88/Domain: fibronectin type II repeat homology <FB2>
F:98-130/Domain: EGF homology <EGL>
F:135-170/Domain: fibronectin type I repeat homology <IF1>
F:178-209/Domain: EGF homology <EG2>
F:217-295/Domain: kringle homology <KRK>
F:298-356/Region: proline-rich
F:354-362,373-615/Product: coagulation factor XIIa, beta form #status experimentally confirmed
F:373-609/Domain: trypsin homology <TRY>
F:98-110,104-119,121-130,135-163,167-170,178-189,183-198,200-209,217-295,238-266-
F:105/binding site: carbohydate (Thr) (covalent) #status experimental
F:249,433/binding site: carbohydate (Asn) (covalent) #status experimental
F:299,305,328,332,337/binding site: carbohydate (Thr) (covalent) #status predicted
F:308/binding site: carbohydate (Ser) (covalent) #status predicted
F:412,461,563/Active site: His, Asp, Ser #status predicted

Query Match	30.2%	Score 681	DB 1	Length 615
Best Local Similarity	34.1%	Pred. No. 3.8e-45		
Matches 152	Conservative 58	Mismatches 158	Indels 78	Gaps 10
QY	13	CLNGGTCVSNKYFSNIHWCNCPKFGGCHCEIDSKTCYEGNGHFFYRGKASTDTMGRPCL 72		
DB	183	CLHGRCLE---VEGHRICHCPVGYTGPFCDVDTKASCYDGRGLSYRGLARTLLSGAPCQ 235		
QY	73	PMSATVLCQTY-HAHRSDALQLGLKHNKYNCRPNDRRRPWCYVQGLKPLVQECMVHDC 131		
DB	240	PWAS---EATYRNVTAEGANWGLGGHAFRCRNPNDIRFCWFLVNRDLRSWEYCYDLAQC 295		
QY	132	-----EATYRNVTAEGANWGLGGHAFRCRNPNDIRFCWFLVNRDLRSWEYCYDLAQC 295		
DB	296	QITPQAAPTVPVSLHVLPLMPAQAPKPKQPTTRTPPQSQTGALPAKBEQPPSLTRNG 355		
QY	137	KTCGCG---KTLPREFKIIGBEFTTIENQPFWAIYRRHRGGSVTVVCGSLISPCWVLIS 192		
DB	356	PLSCGQRUKLSLSSMTRVYGLVARGAHPYIAALWYGH-FCAGSLIAPCWVLIT 405		
QY	194	ATHCFTDIPKXEDYIIVLGRSLKNSNTGEMKMFVENLILHKDYSDATLAHNDIALKKI 253		

F;1-34/Domain: signal sequence #status predicted <SIG>
F;108-148/Domain: fibronectin type II repeat homology <IF2>
F;164-197/Domain: EGF homology <EG1>
F;202-237/Domain: fibronectin-type I repeat homology <FI1>
F;245-278/Domain: EGF homology <EG2>
F;286-367/Domain: kringle homology <KRG>
F;373-407/Product: hepatocyte growth factor activator light chain #status expert
F;408-655/Product: hepatocyte growth factor activator heavy chain #status expert
F;408-641/Domain: trypsin homology <TRY>
F;40, 48, 290, 468, 492, 546/Binding site: carbohydurate (Asn) (covalent) #status pre
F;164-175, 169-185, 188-197, 202-230, 228-237, 245-256, 250-267, 269-278, 286-367, 307-3
F;164-175, 169-185, 188-197, 202-230, 228-237, 245-256, 250-267, 269-278, 286-367, 307-3
F;447, 497, 598/Active site: His, Asp, Ser #status predicted

Query Match 32.2%; Score 726.5; DB 1; Length 655;
Best Local Similarity 37.2%; Pred. No. 1.2e+48;
Matches 158; Conservative 53; Mismatches 161; Indels 53; Gaps 9;

QY 5 HQVPSNCDCLNGGTCVSNKYFSNIHW-----CNCPKPGGOHCEIDSKTKCYEGRGH 56
DB 242 HTACLSPCLNGGTC-----HLIVATGTVCACPPGPAGRLCNIEPBERCELNGT 292

QY 57 FYRKGASTDTMRPCLPWNSATVLQQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCV-- 115.
DB 293 GYRGVASTSASGLSCLAWNDDLYQEIRVDSVGAAALLGLPHAYCRPNPDNDERPMCVVV 352

QY 116 -----OVGIKLPLVEQCWHDCADKLKFQGCK-----TLRPFKI 151
DB 353 KDALSWEYCRLBACESLTVQSFDLLATLEPASPERQ--ACGRHKKTFTLRPR--I 408

QY 152 IGGEFTTIENQPWFIAIYYRHRRGGSVTVCGSGSLISPCWVISATHCFDIYPKKEDIYVL 211
DB 409 IGGSSSLPGSHPLAIIY---IGDS---FCAGSLVHTCWVVSAAHCFSHPERDVSVVU 462

QY 212 GRRLSNSTOGEKKFEVENILIKDYGSADTLAHNDLTALKIRSKEGCAPSRTIQITIC 271
DB 463 QOHFFNRKTTDVOTFGIEKIPIPTLVYSVFNPDSH-DLVLIELKKKGRCATRSEQVPVIC 521

QY 272 LPSMYNDPFQTSCETIFFGENSTDYLPEQLKMVTVKLSIRECCOPHYGYGVSTVMX 331
DB 522 LPREGSTFPAGHKCIAGWGHLDENVGYSSSLREALVELADVADHKCSSPEVVGDISNM 581

QY 332 LCRAADPOWKTDSCGDGSGFLVCSLOQRMTLTGTIVSWGRGCCALKDKDPGYTRVSHFLPW 391
DB 582 LCARGYDFCKDACQDGSGGPLACEKNVAVLYGISWDGCGRLHGKPGVYRVANYVDMVI 641

QY 392 RSHTK 396
DB 642 NDRI R 646

RESULT 17
S28941
coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)
N/Alternate names: Hageman factor
C/Species: Cavia porcellus (guinea pig)
C/Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 21-Jan-2000
C/Acession: S28941
C/Samba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.;
Biochim. Biophys. Acta 1159, 113-121, 1992
A>Title: Primary structure of guinea-pig Hageman factor: sequence around the cl
A/Reference number: S28941; MUJD:93003367; PMID:1390917
A/Acession: S28941
A>Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-603 <SEM>
A/Cross-references: EMBL:X68615; NID:g49578; PID:Caa48600.1; PID:g49579
C/Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat
C/Keywords: hydrolases; serine proteinase
F;46-87/Domain: fibronectin type ii repeat homology <IF2>
F;134-169/Domain: fibronectin type i repeat homology <FI1>
F;177-208/Domain: EGF homology <EGF>
F;216-294/Domain: kringle homology <KRG>
F;359-597/Domain: trypsin homology <TRY>

C:superfamily: tissue plasminogen activator; EGF homology; fibrinectin type I
C:keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-29/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <WAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibrinectin type I repeat homology <IF1>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <KR1>
F:213-294/Domain: kringle homology <KR2>
F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F:309-553/Domain: tryptsin homology <TRY>
F:38-68-66-75-93-94-88-105-107-116-124-205-145-187-176-200-213-294-234-276-265
F:149-484/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 36.8%; Score 829.5; DB 1; Length 559;
Best Local Similarity 35.7%; Pred. No. 1e-56;
Matches 176; Conservative 63; Mismatches 149; Indels 105; Gaps 9

QY 3 ELHQP-----SNCNLNGTCSNKFNIHWCNCPKFGQGHCEIDSKTKCYENGHPY 58
DB 74 QCHSVFVRSCEPRCFNGSGTQQALYDFD-VCQCFDGFVGKRCIDIDTRATCFEEGGITY 132
QY 59 RKGASTDWMRPCLPWNASATVLCQTYHAHRSDALQGLGKENVCHNPNNRRPWCYQVG 118
DB 133 RGTWSTAEGABECINWNSVLSLKPYNARRPNAIKLGLNENYCNPPDRDLKPMCYVFK 192
QY 119 LKPLQECWVHDCADGKLK----- 137
DB 193 GKYTFEFCSTPACPKGEDCVVGKGVYTRGTHSLTTSQASCLPWNISVLMGKSYTAVRT 252
QY 138 -----PQCG-QKTLPKPIIGG 154
DB 253 NSQALGLARHNYCRNPDGADPWCVMKDKRLTWECYDMSPCSTGLGQYKRPQPRIKGG 312
QY 155 EFTTIENQWFAAII-RRHRGSGVTVYCGGLSIPCWISATHCFTIDYPKEDYTVYLR 213
DB 313 LYTDTITSPWQAAIIVFNKRSPGRFLCGGVLISSCWLSAAHCFLERPPNHLKVVLR 372
QY 214 SRLNNTQGMKFEVENILHKVYSADTLAHHNDIALIKIRSEGRCAQPSRTIOTICLP 273
DB 373 TVRVVPGEEQTFEIKYIVHEEPDDDT--YDNDIALQLRSQKCAQSSSVGTACLP 430
QY 274 SMYNDPQF---GTSCITGFKENSTDYLYPEQLKMTWKLISHRECQPHYGVSEVTT 329
DB 431 ----DPNLQLPDWTECELSGYGKEASPPFSDRLKEAHVLYPSSRCTSQLFNKVTIN 486
QY 330 KMLCAADP-----QWKTSQCGDSGGPLVCSLQGRMTITGVISWGRGALKDKPGYTRV 384
DB 487 NMLCAGDTRSGNQDLHDACQDSGGPLVCMINKQMTLTGTIISWLGCGQKDPVGYTKV 546
QY 385 SHELFWIRSHTKK 397
DB 547 TNYLDWIHNNMKQ 559

RESULT 15
JS0600
t-plasminogen activator (BC 3.4.21.68) gamma precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0600
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alago
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampi
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0600
A:Molecule type: mRNA
A:Residues: 1-394 <KRA>
A:Cross-references: GB:M63990; NID:g166078; PIDN:AAA31595.1; PID:g166079

F:311-556/Domain: trypsin homology <TRY>
F:41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203,215-296,236-278,267-291,299-4
F:152,483/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:219/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F:310-311/Cleavage site: Arg-Ile (plasma, trypsin) #status experimental
F:357,406/Active site: His, Asp #status predicted
F:513/Active site: Ser #status experimental

Query Match 38.0%; Score 858.5; DB 1; Length 562;
Best Local Similarity 37.5%; Pred. No. 5,7e-59;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
Db 77 QCHSVPVKSCSEPCFCNGTCQOALYFSDP-VCOCPEGFAGKCEIDTRATCEDOGISY 135
59 RGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQG 118
136 RGTWSTAEGAECTNWNSSALAQPKYSGRRPDAIRLGLGNHNYCRNPDRDRSKPCWCVFKA 195
119 LKPLVQECVHDCADG----- 134
196 KYSSEFCSTPACSEGNSDCVFNGNSAYRGTHSITSGASCLPWNMILGKYVTQNP 255
135 -----KLKFO-----CG-OKTLRPRFKIIGGE 155
256 AQAALGLGKHNYCRNP DGDGAKPWCHLVKNRLTW EYCDVPSCSTCGLRQYQPOFRIKGG 315
156 FTTIENQWPAIYRRH-RGSGVTYVCGGSLISPCWISATHCFIDYPKEDYIVILGRS 214
316 PADTASHPWQAIAIFAKRRSPGERFLCGGILISSCWLSAAHCFERPPHHITVLIGRT 375
215 RLNSNTQCEMKFEVENILHKDYGADTLAHNDIALKIRSKGRCACQAPRTIQTICLPS 274
376 YRVVPGEEQKFEVEKIVHKEFDDT--YDNDIALQLKSDSRCAQESSVVRTCLPP 433
275 MYNDPQFTSCIEITGPKENSTDLYPEOLKMTVKLISHRECQPHYSEVTTMLCA 334
434 ADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSRCSQSHLARTVTDNMLCA 493
335 AD-----POWKT-DSCQDSSGGLVCSLQGRMTLTGIVSGRGKALCKDKPGVYTRVSHFL 388
494 GDRSGGQANLHDACQDSSGGLVCLNDGRMTLVGLIISWGLGCGQKDXVGVYTKVNYL 553
389 PWIRSHTK 396
554 DWIRDNR 561

RESULT 12
A35029
t-plasminogen activator (EC 3.4.21.68) precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A35029; A31597
R: Feng, P.; Ohlsson, M.; Ny, T.
J. Biol. Chem. 265, 2022-2027, 1990
A:Title: The structure of the TATA-less rat tissue-type plasminogen activator gene. Spec
A:Reference number: A35029; MUID: 90130448; PMID: 2105315
A:Accession: A35029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-559 <P>
A:Cross-references: GB:M31197; NID: g207429; PIDN: AAA42261.1; PID: g207431; GB: J05226
R: Ny, T.; Leonardsson, G.; Heueh, A.J.W.
DNA 7, 671-677, 1988
A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activat
A:Reference number: A31597; MUID: 89170114; PMID: 3148445
A:Accession: A31597
A:Molecule type: mRNA
A:Residues: 1-379, 'K', 381-559 <N>
A:Cross-references: GB:M23697; NID: g530159; PID: g530160
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom

C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-29/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:38-75/Domain: fibronectin type I repeat homology <ACH>
F:93-116/Domain: EGF homology <EGF>
F:124-203/Domain: kringle homology <K1>
F:213-294/Domain: kringle homology <K2>
F:309-553/Domain: trypsin homology <TRY>
F:38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-
F:149,48/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Ile (plasma, trypsin) #status predicted
F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 37.3%; Score 841.5; DB 1; Length 559;
Best Local Similarity 36.5%; Pred. No. 1.2e-57;
Matches 180; Conservative 65; Mismatches 143; Indels 105; Gaps 12;
QY 3 ELHQP-----SNCCDCLNGTCVSNKYSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFY 58
Db 74 QCHSVPVKSCSEPCFCNGTCQOALYFSDP-VCOCPEGFAGKCEIDTRATCEDOGISY 132
QY 59 RGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYV-QV 117
Db 133 RGTWSTAEGAECTNWNSSALAQPKYSGRRPDAIRLGLGNHNYCRNPDRDRSKPCWCVFKA 192
QY 118 GL-----KPLVQECMV----- 128
Db 193 GKYTEFCSTPACPKPTEDCVGKGYRGTHFTTSKASCLPWNMILGKYVTAWPA 252
QY 129 -----HDCADGKLKFO-----CGOKTLR-PRFKIIGG 154
Db 253 NSQAALGLGKHNYCRNP DGDGAKPWCHLVKNRLTW EYCDVPSCSTCGLRQYQPOFRIKGG 312
QY 155 ETTIENQWPAIYRRH-RGSGVTYVCGGSLISPCWISATHCFIDYPKEDYIVILGR 213
Db 313 LFTDITSHPWQAIAIFAKRRSPGERFLCGGILISSCWLSAAHCFERPPHHLKVILGR 372
QY 214 RLNSNTQCEMKFEVENILHKDYGADTLAHNDIALKIRSKGRCACQAPRTIQTICL 273
Db 373 YRVVPGEEQKFEVEKIVHKEFDDT--YDNDIALQLKSDSRCAQESSVGTACL 430
QY 274 SMYNDPQF-----GTSCEITGPKENSTDLYPEOLKMTVKLISHRECQPHYSEVTT 329
Db 431 ----DPDVLQPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSRCSQSHLFNKTITS 486
QY 330 KMLCAADP-----QWKTDSCQDSSGGLVCSLQGRMTLTGIVSGRGKALCKDKPGVYTRV 384
Db 487 NMLCAGDTGTGNQDVHDACQDSSGGLVCLNDGRMTLVGLIISWGLGCGQKDXVPGIYTKV 546
QY 385 SHFLPWIRSHTK 397
Db 547 TNYLAWIQDNMKQ 559

RESULT 13
JS0597
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0597
R: Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampir
A:Reference number: JS0597; MUID: 92039036; PMID: 1937019
A:Accession: JS0597
A:Molecule type: mRNA
A:Residues: 1-477 <RA>
A:Cross-references: GB:M63987; NID: g166070; PIDN: AAA31591.1; PID: g166071
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I r

RESULT 11

UKHT

t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
 N/Alternate names: t-PA; tissue plasminogen activator
 C/Species: Homo sapiens (man)
 C/Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 08-Dec-2000
 C/Accession: A94004; A23529; J0562; A93293; S02125; A93543; A93522; A54645; 160
 R/RY, T.; Elgin, F.; Lund, B.
 Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
 A/Title: The structure of the human tissue-type plasminogen activator gene: correlation
 A/Reference number: A94004; MUID:84298137; PMID:6089198
 A/Accession: A94004
 A/Molecule type: DNA
 A/Residues: 1-562 <NYT>
 A/Cross-references: GB:J00141
 A/Note: The codon given for residue 93 (ACC) is inconsistent with the authors' translation
 R/Friener Degen, S.J.; Rajput, B.; Reich, E.
 J. Biol. Chem. 261, 6972-6985, 1986
 A/Title: The human tissue plasminogen activator gene.
 A/Reference number: A23529; MUID:86196143; PMID:3009482
 A/Accession: A23529
 A/Molecule type: DNA
 A/Residues: 1-562 <DEG>
 A/Cross-references: GB:K03021; NID:G339817; PIDN:AAA98809.1; PID:G339818
 R/Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuoka, S.; Higashio, K.
 Agric. Biol. Chem. 55, 1225-1232, 1991
 A/Title: Purification and characterization of tissue plasminogen activator secreted by H
 A/Reference number: J0562; MUID:91291340; PMID:1368681
 A/Accession: J0562
 A/Molecule type: mRNA
 A/Residues: 31-562 <ITA>
 A/Cross-references: DDBJ:D01096; NID:G220128; PIDN:BAA00881.1; PID:G441174
 A/Experimental source: embryonic lung fibroblast IMR-90 cells
 A/Note: Part of this sequence, including the amino end of the mature protein, was confir
 R/Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vohar, G.A.; Ward, C.A.; Bennett
 Nature 301, 214-221, 1983
 A/Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
 A/Reference number: A93293; MUID:83115262; PMID:6337343
 A/Accession: A93293
 A/Molecule type: mRNA
 A/Residues: 1-562 <PEN>
 A/Cross-references: GB:J00141
 A/Experimental source: melanoma cells
 R/Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
 Nucleic Acids Res. 16, 5695, 1988
 A/Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
 A/Reference number: S02125; MUID:88262579; PMID:3133640
 A/Accession: S02125
 A/Status: translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-562 <SAS>
 A/Cross-references: EMBL:X07933; NID:G37243; PIDN:CAA30302.1; PID:G37244
 R/Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma
 FEBS Lett. 199, 145-149, 1985
 A/Title: Expression in *Escherichia coli* of finger-domain lacking tissue-type plasminogen
 A/Reference number: A93343; MUID:85285620; PMID:3996853
 A/Accession: A93343
 A/Molecule type: mRNA
 A/Residues: 1-38, 'G', 435-562 <KAG>
 A/Experimental source: Detroit 562 cells; ATCC 138
 R/Selund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
 Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
 A/Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activ
 A/Reference number: A93951; MUID:83169656; PMID:6572897
 A/Accession: A93951
 A/Molecule type: mRNA
 A/Residues: 251-358 <EDL>
 A/Experimental source: melanoma cells
 R/Fohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.
 Biochemistry 23, 3701-3707, 1984

A/Title: Tissue plasminogen activator: peptide analyses confirm an indirectly
 differences.
 A/Reference number: A90488; MUID:85000468; PMID:6433976
 A/Contents: annotation; melanoma cells, partial sequence of residues 36-562, a
 R/Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.
 FEBS Lett. 168, 29-32, 1984
 A/Title: Differences between uterine and melanoma forms of tissue plasminogen
 A/Reference number: A91322; MUID:84158956; PMID:6538514
 A/Accession: A91322
 A/Molecule type: protein
 A/Residues: 33-45,311-320 <POH>
 A/Experimental source: uterus
 A/Note: In the uterus, cleavage of the activation peptide may also occur after
 R/van Zonneveld, A.J.; Veerman, H.; Pannetkoek, H.
 J. Biol. Chem. 261, 14214-14218, 1986
 A/Reference number: A37567; MUID:87033611; PMID:3021732
 A/Contents: annotation; fibrin binding site
 R/Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, F.
 EMBO J. 5, 3525-3530, 1986
 A/Title: Involvement of finger domain and kringle 2 domain of tissue-type plas
 A/Reference number: A37568; MUID:87161761; PMID:3030730
 A/Contents: annotation; fibrin binding site
 R/Dodd, I.; Nunn, B.; Robinson, J.H.
 Thromb. Haemost. 59, 523-528, 1988
 A/Title: Isolation, identification and pharmacokinetic properties of human tis
 A/Reference number: A60902; MUID:89044681; PMID:3142086
 A/Contents: annotation; novel forms of expressed recombinant t-PA
 R/Harris, T.J.R.; Patel, T.; Marsden, F.A.O.; Little, S.; Emrage, J.S.; Opdena
 Mol. Biol. Med. 3, 279-292, 1986
 A/Title: Cloning of cDNA coding for human tissue-type plasminogen activator at
 A/Reference number: A54645; MUID:86284200; PMID:3090401
 A/Accession: A54645
 A/Molecule type: mRNA
 A/Residues: 1-562 <HAR>
 A/Cross-references: GB:M15518; NID:G190031; PIDN:AAA60111.1; PID:G190032
 A/Note: Parts of this sequence were confirmed by peptide sequencing
 R/Reddy, V.B.; Garramone, A.J.; Sasak, H.; Wei, C.
 DNA 6, 461-472, 1987
 A/Title: Expression of human uterine tissue-type plasminogen activator in mous
 A/Reference number: I60110; MUID:88054470; PMID:2824147
 A/Accession: I60110
 A/Status: translated from GB/EMBL/DDBJ
 A/Molecule type: mRNA
 A/Residues: 1-562 <RES>
 A/Cross-references: GB:M18182; NID:G340176; PIDN:AAA36800.1; PID:G340177
 R/Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, D.
 J. Biol. Chem. 260, 11223-11230, 1985
 A/Title: Isolation and characterization of the human tissue-type plasminogen a
 A/Reference number: I55232; MUID:85289338; PMID:3161893
 A/Accession: I55232
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: DNA
 A/Residues: 1-36 <RE2>
 A/Cross-references: GB:M11890; NID:G339837; PIDN:AAA61213.1; PID:G339839
 C/Comment: Cleavage by plasmin or trypsin produces two chains held together by
 C/Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val
 C/Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type
 C/Genetics:
 A/Gene: GDB:PLAT
 A/Cross-references: GDB:119496; OMIM:173370
 A/Map position: 8p12-8p12
 A/Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1;
 C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I
 C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine pr
 F/1-23/Domain: signal sequence #status predicted <SIG>
 F/24-32/Domain: propeptide #status predicted <PRO>
 F/33-562/Product: t-plasminogen activator #status experimental <MAT>
 F/33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
 F/46-119/Domain: fibronectin type I repeat homology <1F1>
 F/81-78/Domain: EGF homology <EGF>
 F/127-208/Domain: kringle homology <KR1>
 F/215-296/Domain: kringle homology <KR2>
 F/311-562/Product: t-plasminogen activator chain B #status experimental <BCH>

5/3;
 eat
 inase

```
QY 59 RKASTDWGRCLPWNATVLQOQTYHAHRSDALQLGLGKNYCRNPNRRRPPWCYVQG 118
Db 137 RGTWSTSSGAQCINWNSNLLTRTYNGRRSDAITLGLGNHYCRNPNRRRPPWCYVKA 196
QY 119 LKPLVQECMVHDCADGKLFQCG-QKTLRPRFKIIGGEFTTIENQPFAPAAIYRHRGGS- 176
Db 197 SKFILEFCSVPVCS----KATCGLRKYKEPQLHSTGGGLFTDITSHPWQAALFAQNRSSG 252
QY 177 VTYVCGGSLISPCWISATHCFID-YPKEDYIVYVGRSLNNTQGMKFEVENLILHK 235
Db 253 ERFLOGGILISSCWLTAAHCFQERYPPQHLRV-LGRTRYVKPKGEQTEFEVEKCIVHE 311
QY 236 DYSADTLAHDNDIALKIRSKEGRCQAQPSRTIQTCLPSMYNDPQFGTSCEITGFGKENS 295
Db 312 EBDTDT--YNDIALQLKSGSPQCAQESDSVRAICLPEANLQLPDWTCELSYGKHS 369
QY 296 TDLYPEQIKMTVVKLISHRSCQPHYYGSEVTTKMLCAADPQWKT-----DSCQSDSG 349
Db 370 SSPFFYSEQLKEGHVLYPSSRCTSKFLFNKVTYTNMLCAGDTRSGEIPNVHDAQCQSDG 429
QY 350 GPLVCSLQGRMTLTGIVSWGRCALKDKPGVYVTRVSHFLPWRSHTK 396
Db 430 GPLVCRNDNHMTLLGIISWVGCGEKPFGVYTKVNTYLGWIRDNMR 476

RESULT 9
A34369
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
C:Species: Megaderma lyra
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A34369
R:Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob
J. Biol. Chem. 264, 17947-17952, 1989
A:Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin
A:Reference number: A34369; MUID:90036867; PMID:2509450
A:Accession: A34369
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-477 <GAR>
A:Cross-references: GB:J05082; NID:G166080; PIDN:AAA31596.1; PID:G166081
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <IFA>
F:128-209/Domain: EGF homology <EGF>
F:226-471/Domain: kringle homology <KRG>
F:42-72,70-98,87-98,92-109,111-120,128-128 <TRY>
F:42-72,70-98,87-98,92-109,111-120,128-128 <TRY>
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 38.4%; Score 867.5; DB 1; Length 477;
Best Local Similarity 43.7%; Pred. No. 9.5e-60;
Matches 178; Conservative 60; Mismatches 148; Indels 21; Gaps 9;

QY 3 ELHQVP----SNCDCLNGTCVSNKYFNSNIHWCNCPKFGGHCIDSKTCYEGNGHFY 58
Db 78 QCHTVPVKSCSELRCFNGTGWQAASFSDF-VCQCPKGYTGKQCEVDTHATCYKQGVY 136
QY 59 RKASTDWGRCLPWNATVLQOQTYHAHRSDALQLGLGKNYCRNPNRRRPPWCYVQG 118
Db 137 RGTWSTSSGAQCINWNSNLLTRTYNGRRSDAITLGLGNHYCRNPNRRRPPWCYVKA 196
QY 119 LKPLVQECMVHDCADGKLFQCG-QKTLRPRFKIIGGEFTTIENQPFAPAAIYRHRGGS- 176
Db 197 SKFILEFCSVPVCS----KATCGLRKYKEPQLHSTGGGLFTDITSHPWQAALFAQNRSSG 252
QY 177 VTYVCGGSLISPCWISATHCFID-YPKEDYIVYVGRSLNNTQGMKFEVENLILHK 235
Db 253 ERFLOGGILISSCWLTAAHCFQERYPPQHLRV-LGRTRYVKPKGEQTEFEVEKCIVHE 311
```

```
QY 236 DYSADTLAHDNDIALKIRSKEGRCQAQPSRTIQTCLPSMYNDPQFGTSCEITGFGKENS 295
Db 312 EBDTDT--YNDIALQLKSGSPQCAQESDSVRAICLPEANLQLPDWTCELSYGKHS 369
QY 296 TDLYPEQIKMTVVKLISHRSCQPHYYGSEVTTKMLCAADPQWKT-----DSCQSDSG 349
Db 370 SSPFFYSEQLKEGHVLYPSSRCTSKFLFNKVTYTNMLCAGDTRSGEIPNVHDAQCQSDG 429
QY 350 GPLVCSLQGRMTLTGIVSWGRCALKDKPGVYVTRVSHFLPWRSHTK 396
Db 430 GPLVCRNDNHMTLLGIISWVGCGEKPFGVYTKVNTYLGWIRDNMR 476

RESULT 10
JS0599
t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0599
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boisdol, W.; Bringmann, P.; Alagon
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0599
A:Molecule type: mRNA
A:Residues: 1-431 <KRA>
A:Cross-references: GB:M63989; NID:G166076; PIDN:AAA31594.1; PID:G166077
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I r
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-431/Product: plasminogen activator beta #status predicted <PLA>
F:41-74/Domain: EGF homology <EGF>
F:82-163/Domain: kringle homology <KRG>
F:180-425/Domain: trypsin homology <TRY>
F:41-52,46-63,65-74,82-163,103-135,134-158,168-299,211-227,219-288,313-388/Disu
F:439,352/Binding site: carbohydrate (Aen) (covalent) #status predicted
F:179-180/Cleavage site: His-Ser (plasmin) #status predicted
F:226,275,382/Active site: His, Asp, Ser #status predicted
F:345-361,378-406/Disulfide bonds: #status predicted

Query Match 38.2%; Score 862.5; DB 2; Length 431;
Best Local Similarity 44.1%; Pred. No. 2.1e-59;
Matches 175; Conservative 59; Mismatches 146; Indels 17; Gaps 8;

QY 9 SNCDCLNGTCVSNKYFNSNIHWCNCPKFGGHCIDSKTCYEGNGHFYRGKASTDTWG 68
Db 42 SELRCFNGTGWQAASFSDF-VCQCPKGYTGKQCEVDTHATCYKQGVYRGKASTDTWG 100
QY 69 RCLPWNATVLQOQTYHAHRSDALQLGLGKNYCRNPNRRRPPWCYVQVGLKPLVQECMV 128
Db 101 AQCNWNSNLLTRTYNGRRSDAITLGLGNHYCRNPNRRRPPWCYVQVGLKPLVQECMV 160
QY 129 HDCADGKLFQCG-QKTLRPRFKIIGGEFTTIENQPFAPAAIYRHRGGS-VTYVCGGSLI 186
Db 161 PVCS----KATCGLRKYKEPQLHSTGGGLFTDITSHPWQAALFAQNRSSGERFLCGGILI 216
QY 187 SPQWISATHCFID-YPKEDYIVYVGRSLNNTQGMKFEVENLILHKYSATLAHH 245
Db 217 SSCWLTAAHCFQERYPPQHLRV-LGRTRYVKPKGEQTEFEVEKCIIEEEDDTT--YN 273
QY 246 NDIALKIRSKEGRCQAQPSRTIQTCLPSMYNDPQFGTSCEITGFGKENSVDLYPEQIK 305
Db 274 NDIALQLKSGSPQCAQESDSVRAICLPEANLQLPDWTCELSYGKHSPPFSEQLK 333
QY 306 MTVVKLISHRSCQPHYYGSEVTTKMLCAADPQWKT-----DSCQSDSGGPLVCSLQGR 359
Db 334 EGHVLYPSSRCTSKFLFNKVTYTNMLCAGDTRSGEIPNVHDAQCQSDSGGPLVCMNDNH 393
QY 360 MTLTGIVSWGRCALKDKPGVYVTRVSHFLPWRSHTK 396
Db 394 MTLGLIISWVGCGEKPFGVYTKVNTYLGWIRDNMR 430
```

A;Accession: A24615
A;Molecule type: mRNA
A;Residues: 1-433 <BEL>
A;Cross-references: GB:X02389; NID:G55127; PIDN:CAA26231.1; PID:G55128
C;Genetics:
A;Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;32-63/Domain: EGF homology <EGF>
F;71-132/Domain: kringle homology <KR>
F;180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F;180-421/Domain: trypsin homology <TRY>
F;169-301.211-227,219-230,315-384,347-363,374-402/Disulfide bonds: #status predicted
F;226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 72.1%; Score 1626.5; DB 1; Length 433;
Best Local Similarity 70.2%; Pred. No. 1e-118;
Matches 283; Conservative 50; Mismatches 61; Indels 9; Gaps 2;

QY 9 SNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHFYRGKASTDTMG 68
DB 30 SNGCQNGGVCSVKYFSRIRRCSPKFGQHCEDASKTCYHNGDSYRGKANTDTKG 89
QY 69 RPLPWSATVLO-QTYHAHRSALQGLGKHNYCRPNRRPWCYVQVGLKPLVQECMV 128
DB 90 RPLANWAPVLOKPYNAHDPDAISLGLGRHNYCRPNQKREWCYVQGLRQVQECMV 149
QY 129 HCDADGKXK-----FQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGG-VTY 179
DB 150 HCDLSLKKPSSVDQCGQKALRPRFKIIGGEFTTIENQWFAAIYRHRGGGPPSF 209
QY 180 VCGSLSPCWISATHCFIDYPRKEDIYVYLGSRNSNTQGMKFEVENLILHKDYSA 239
DB 210 KCGSLSPCWASAAHCFIQLPKENYVYVYLGSKSSYNPCKMFEVEQLTLHYYRE 269
QY 240 DTLAHNDIALKIRSKGRCAOPSRITQICLPSMYNDPQFCTSCITGFGKENSVDYL 299
DB 270 DSLAYNDIALKIRSKGRCAOPSRITQICLPSMYNDPQFCTSCITGFGKENSVDYL 329
QY 300 YPEQLQVTVKLSHRECOQPHYGVSEVTKMLCAADPQWKTDCQDGSGLPVCLOQR 359
DB 330 YPKNLKMSVVKLSHRECOQPHYGVSEVTKMLCAADPQWKTDCQDGSGLPVCLOQR 389
QY 360 MTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHYKENGCLA 402
DB 390 PTLGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHYKENGCLA 432

RESULT 7
A35005
U-Plasminogen activator (EC 3.4.21.73) precursor - chicken
N;Alternate names: uPA
C;Species: Gallus gallus (chicken)
C;Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
C;Accession: A35005
R;Leslie, N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.
J. Biol. Chem. 265, 1339-1344, 1990
A;Title: The chicken urokinase-type plasminogen activator gene.
A;Reference number: A35005; MUID:90110185; PMID:2295632
A;Accession: A35005
A;Status: Preliminary
A;Molecule type: mRNA
A;Residues: 1-434 <LES>
A;Cross-references: GB:J05187; NID:G212858; PIDN:AAA49131.1; PID:G212859
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;40-71/Domain: EGF homology <EGF>
F;79-158/Domain: kringle homology <KR>
F;173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>

F;173-416/Domain: trypsin homology <TRY>
F;162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status pre
F;217,272,373/Active site: His, Asp, Ser #status predicted

Query Match 44.2%; Score 997; DB 1; Length 434;
Best Local Similarity 47.8%; Pred. No. 7.6e-70;
Matches 187; Conservative 64; Mismatches 122; Indels 18; Gaps 7

QY 11 CDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHFYRGKASTDTWGRP 70
DB 40 COCLNGGTCITRFRFSQIKRCLCEGGLHCEIDTNSICYSNGGEDYRGNAEDP----G 95
QY 71 CLPWSATVLO-QTYHAHRSALQGLGKHNYCRPNRRPWCYVQVGLKPLVQECMV 127
DB 96 CLYNDHPSVIRWDYHADLKNALQGLGKHNYCRPNRGRSPWCYTK--RRYSIOETPCS 153
QY 128 VHDCAKGKLFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGGVTYVCGSLIS 187
DB 154 TIE-----KCEFTCGQSPSKYFKIVGSGQAEVETQPIAGFINIM-GTDFLCGSLID 208
QY 188 PCWISATHCFID----YPKKEDIYVYLGSRNSNTQGMKFEVENLILHKDYSAADTLA 243
DB 209 PCWVLTAAHCFVNFPTKQPNKSVYKFLGKSLANTDEHEQVFWVDEIISHPDTDTGG 268
QY 244 HHNDIALKIRSKGRCAOPSRITQICLPSMYNDPQFCTSCITGFGKENSVDYLPEQ 303
DB 269 NDNDIALIRIRTAGSQCAVESNYRTVCLPEKNLYDNTWCEIAGYKQNSYDIYAQR 328
QY 304 LKMTVKLISHRECOQPHYGVSEVTKMLCAADPQWKTDCQDGSGLPVCLOQWMTLT 363
DB 329 LMSATVNLISQDDCKNKYDSTRVTDNMCAGDPLWETDACKGSGGPMVCEHNGMTLY 386
QY 364 GIVSWGRGCAKDKPGVYTRVSHFLPWIRSH 394
DB 389 GIVSWGDCAKKNKPGVYTRVTRVYLNMDSN 419

RESULT 8
J50598
U-Plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: J50598
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagc
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampi
A;Reference number: J50597; MUID:9203036; PMID:1937019
A;Accession: J50598
A;Molecule type: mRNA
A;Residues: 1-477 <KRA>
A;Cross-references: GB:M63988; NID:G166074; PIDN:AAA31593.1; PID:G166075
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
F;42-79/Domain: fibronectin type I repeat homology <1PA>
F;87-120/Domain: EGF homology <EGF>
F;128-209/Domain: kringle homology <KR>
F;142-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265
F;185,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;225-226/Cleavage site: His-Ser (plasma) #status predicted
F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 38.5%; Score 868.5; DB 2; Length 477;
Best Local Similarity 43.7%; Pred. No. 8e-60;
Matches 178; Conservative 60; Mismatches 148; Indels 21; Gaps 5

QY 3 ELHQVPP-----SNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDASKTCYEGNGHFY 58
DB 78 QCHTVFVKSCELRCFNGGTCQAASFSDF-VCCPKYTKGQCEVDHANCYKQGVTY 136

A;Molecule type: mRNA
A;Residues: 1-433 <KRA>
A;Cross-references: GB:L03546; NID:G163800; PIDN:AA51419.1; PID:G163801
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-179/Product: plasminogen activator chain A #status predicted <MA1>
F;21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;33-64/Domain: EGF homology <EGF>
F;72-153/Domain: kringle homology <KRG>
F;181-433/Product: plasminogen activator chain B #status predicted <MA2>
F;181-433/Product: plasminogen activator chain B #status predicted <MA2>
F;170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F;226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 76.6%; Score 1728; DB 1; Length 433;
Best Local Similarity 73.8%; Pred. No. 1.4e-126;
Matches 305; Conservative 44; Mismatches 54; Indels 10; Gaps 2;

QY 1 SNEHQV--PNCDCNLNGTCTVSNKYFSNIHWCNPKKFGQGHCEIDKSKTCYEGNGHFY 58
DB 21 SNEVKEGSGESGCLNGKGVTKYFSNIQRCSCPKKFGQGHCEIDKSKTCYQGNHGY 80

QY 59 RGKASTDMGPCLPWNATVLOQTYHAHRSDALQLGLGKHNCRNPNRRPWCYVQVG 118
DB 81 RGNARDLSGRPCLANSDPTVLLKMYHAHRSDALQLGLGKHNCRNPNRRPWCYVQIG 140

QY 119 LKPLVQECMVHDCADGKL-----XPCQCKTLRPRFKIIGSEFTTIENQWPAIYR 170
DB 141 LKQFVQCVQWDCSVGKSPSPREKEBFQCCQKALRPRFKIIGSEFTTIENQWPAIYR 200

QY 171 RHRGGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNQTGEMKPEVEN 230
DB 201 RHRGGSITVLCGSLISPCWVVSATHCFIDHPKXENIVVIGQRLNSDTRGEMQFEVK 260

QY 231 LILHKDYSADTLAHHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCITGF 290
DB 261 LILHEDYSAESLAHHNDIALKIRTSRGCAQPSRSIQTICLPRPEHDAHSRTRCITGF 320

QY 291 GKENSTVLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDCQDSSG 350
DB 321 GKENPSDVRYSDELKMTFVLSVSEVQCPHYGAEVTDKMLCAADPQWETDSCQDSSG 380

QY 351 PLVSLQRMVTLTGIVSGRCALKDKPGVTVRVSHFPLWIRSHKTEENGLAL 403
DB 381 PLVCTIQRLTLTGIVSGRCACAKYKPGVTVRVSKFLPWINTHTRGEINLV 433

RESULT 5
S18932
u-plasminogen activator (EC 3.4.21.73) precursor - rat
N;Alternate names: plasminogen activator, urokinase-type, urinary plasminogen activator
C;Species: Rattus norvegicus (Norway rat)
C;Date: 18-Oct-1989 #sequence revision 10-Feb-1995 #text_change 18-Jun-1999
C;Accession: S24604; I60186; I53472; S18932
R;Rabbani, S.A.
submitted to the EMBL Data Library, April 1992
A;Reference number: S24604
A;Accession: S24604
A;Molecule type: mRNA
A;Residues: 1-15,'H',17-23,'G',25-331,'N',333-432 <RAB>
A;Cross-references: EMBL:X65651; NID:G57456; PIDN:CAA46601.1; PID:G57457
A;Experimental source: tissue kidney
R;Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
Cancer Res. 52, 2489-2496, 1992
A;Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act
A;Reference number: I60186; MUID:92233409; PMID:1568219
A;Accession: I60186
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-432 <RES>
A;Cross-references: EMBL:X63434; NID:G57465; PIDN:CAA45028.1; PID:G57466
A;Experimental source: strain Fischer 344; tissue mammary

R;Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Biasi, F.; Rossi, G.
PEBS Lett. 306, 193-198, 1992
A;Title: The receptor for the plasminogen activator of urokinase type is up-reg
A;Reference number: I53472; MUID:92339549; PMID:I321734
A;Accession: I53472
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 31-62 <RE2>
A;Cross-references: EMBL:X66907; NID:G396200; PIDN:CAA47356.1; PID:G938279
C;Genetics:
A;Gene: uPA
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homo
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-177/Product: urokinase-type plasminogen activator chain A #status predicte
F;31-62/Domain: EGF homology <EGF>
F;70-151/Domain: kringle homology <KRG>
F;179-433/Product: urokinase-type plasminogen activator chain B #status predict
F;179-433/Domain: trypsin homology <TRY>
F;168-300,210-226,218-289,314-383,346-362,373-401/Disulfide bonds: #status pred
F;225,276,377/Active site: His, Asp, Ser #status predicted

Query Match 73.4%; Score 1656.5; DB 1; Length 432;
Best Local Similarity 72.2%; Pred. No. 4.8e-121;
Matches 291; Conservative 45; Mismatches 58; Indels 9; Gaps 2;

QY 9 SNDCCLNGTCTVSNKYFSNIHWCNPKKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMG 68
DB 29 SNGCGQNGVCVSYKYFSIRRCSCPKKFGQGHCEIDKSKTCYEGNGOSYRGKANTDTKG 88

QY 69 RCLPWNATVLOQTYHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLKPLVQECMV 128
DB 89 RCLPWNATVLOQTYHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLKPLVQECMV 148

QY 129 HDCADGKL-----FQCQCKTLRPRFKIIGSEFTTIENQWPAIYRHRGGS-VTY 179
DB 149 QCCLSKXKSPSTVDCQCFQCCQKALRPRFKIIGSEFTTIENQWPAIYRHRGGSPPSF 208

QY 180 VCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNQTGEMKPEVENLILHKDYS 239
DB 209 KCGSLISPCWVVSATHCFVNPQKKEVYVILGSKNSYNPGEKMFVEQLIHEDFS 268

QY 240 DTLAHHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCITGFGENSTDY 299
DB 269 ETAFHNDIALKIRTSRGCAQPSRTIQTICLPRFGDAPFGSDCEITFGQESATDYF 328

QY 300 YPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDCQDSSGGLVCSLOGR 359
DB 329 YPKDLKMSVVKIISHQCKQPHYGVSEVTTKMLCAADPQWKTDCQDSSGGLVCSLOGR 388

QY 360 MTLTGIVSGRCALKDKPGVTVRVSHFPLWIRSHKTEENGLA 402
DB 389 PTLTGIVSGRCALKDKPGVTVRVSHFPLWIRSHKTEENGLA 431

RESULT 6
UKMS
u-plasminogen activator (EC 3.4.21.73) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C;Accession: A29420; A24615
R;Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
Biochemistry 26, 8270-8279, 1987
A;Title: The murine urokinase-type plasminogen activator gene.
A;Reference number: A29420; MUID:86163489; PMID:2831940
A;Accession: A29420
A;Molecule type: DNA
A;Residues: 1-433 <DEG>
A;Cross-references: GB:M17922; NID:G202296; PIDN:AAA40539.1; PID:G202297
R;Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.
Eur. J. Biochem. 148, 225-232, 1985
A;Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse
A;Reference number: A24615; MUID:85179474; PMID:2985383

QY 353 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGGLAL 403
Db 381 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGGLAL 431

RESULT 2

u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
C:Accession: S14687; S08651
R:Altus, Y.P.T.; Wang, T.W.; Clowes, A.W.
Nucleic Acids Res. 18, 3411, 1990
A:Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen
A:Reference number: S14687; MUID:90287734; PMID:2113276
A:Accession: S14687
A:Molecule type: mRNA
A:Residues: 1-433 <AUY>
A:Cross-references: EMBL:X51935; NID:938130; PIDN:CAA36200.1; PID:938131
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-176/Product: plasminogen activator chain A #status predicted <ACH>
F:30-61/Domain: EGF homology <EGF>
F:69-150/Domain: kringle homology <KRG>
F:178-433/Product: plasminogen activator chain B #status predicted <BCH>
F:178-433/Domain: trypsin homology <TRY>
F:167-298, 208-224, 216-287, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
F:223, 274, 378/Active site: His, Asp, Ser #status predicted
F:324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 92.6%; Score 2090; DB 1; Length 433;
Best Local Similarity 90.6%; Pred. No. 1.1e-154;
Matches 375; Conservative 17; Mismatches 10; Indels 12; Gaps 3;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 21 SREL-QVPSDCGLNGGTCVSNKYFSNIHMCNCPKFGGQHCIDKSKTCYEGNGHYRG 79
QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHYCRNPDRRPPWCYVQVGLK 120
Db 80 KASTDTMGRSCLANWNSATVLTQTYHAHRSALQGLGKHYCRNPDRRPPWCYVQVGLK 139
QY 121 PLVQECMVHDCADGK-----LKFCGGQKTLRPRFKIIGGFTTIENOPWFAALYRRH 172
Db 140 QRVOECMVHDCADGKSSPEELQFCGGQKTLRPRFKIIGGFTTIENOPWFAALYRRH 199
QY 173 RGGSVTVVCGGSLISPCWVWSATHCFIDYPKKEDYIYVLSGRSLNSNTQGMKFEVENLI 232
Db 200 RGGSVTVVCGGSLISPCWVWSATHCFIDYPKKEDYIYVLSGRSLNSNTQGMKFEVENLI 259
QY 233 LHKDYSADTLAHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQ---FGTSCEITG 289
Db 260 LHEDYSADTLAHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPDPFPFGTSCEITG 319
QY 290 FGKENSVDLYPEQLKMTVVKLIHSRECCQPHYVSGVTTKMLCAADPQWTKTSQCGDSG 349
Db 320 FGKENSVDLYPEQLKMTVVKLVSHQKCCQPHYVSGVTTKMLCAADPQWTKTSQCGDSG 379
QY 350 GPLVCSIQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGGLAL 403
Db 380 GPLVCSIQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGGLAL 433

RESULT 3

u-plasminogen activator (EC 3.4.21.73) precursor - pig
N:Alternate names: uPA
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
C:Accession: A00932
R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
Nucleic Acids Res. 12, 9525-9541, 1984

A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A:Reference number: A00932; MUID:85087954; PMID:6096832
A:Accession: A00932
A:Molecule type: DNA
A:Residues: 1-240, 'H', 242-442 <NAG1>
A:Experimental source: kidney cell line LLC-PK1
R:Nagamine, Y.
submitted to the Protein Sequence Database, December 1986
A:Reference number: A37566
A:Contents: annotation; correction to residue 241
C:Genetics:
F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted
F:190-430/Domain: trypsin homology <TRY>
F:152/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:179-310, 220-236, 228-299, 324-393, 356-372, 383-411/Disulfide bonds: #status predicted
F:235, 286, 387/Active site: His, Asp, Ser #status predicted

Query Match 81.0%; Score 1827.5; DB 1; Length 442;
Best Local Similarity 78.1%; Pred. No. 2.6e-134;
Matches 329; Conservative 32; Mismatches 41; Indels 19; Gaps 2;

QY 1 SNEHQV--PSNCDCLNGGTCVSNKYFSNIHMCNCPKFGGQHCIDKSKTCYEGNGHY 58
Db 21 SHELHQESGASNCGLNGGKCVSYKFSNIQRCSCPKKFGQHCIDKSKTCYEGNGHSY 80
QY 59 RKGASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHYCRNPDRRPPWCYVQV 110
Db 81 RKGANTNTGGRPCLPWNSATVLTQTYHAHRSALQGLGKHYCRNPDRRPPWCYVQV 140
QY 119 LKPLVQECMVHDCADG-----LKFCGGQKTLRPRFKIIGGFTTIEN 161
Db 141 LKQLVQECMVHDCADGSHRPFAYDCKNPFSTPEKVFQCGKALRPRFKIIGGFTTIEN 200
QY 162 QPWFPAIYRRHGGSVTVVCGGSLISPCWVWSATHCFIDYPKKEDYIYVLSGRSLNSNTQ 221
Db 201 QPWFPAIYRRHGGSVTVVCGGSLISPCWVWSATHCFIDYPKKEDYIYVLSGRSLNSNTQ 260
QY 222 GSKFEVENLILHKDYSADTLAHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQF 281
Db 261 GSKFEVENLILHKDYSADTLAHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQF 320
QY 282 GTSCEITGFGKENSVDLYPEQLKMTVVKLIHSRECCQPHYVSGVTTKMLCAADPQWTK 341
Db 321 GASCEIVGFGKEDPSDLYPEQLKMTVVKLVSHRECCQPHYVSGVTTKMLCAADPQWTK 380
QY 342 DSCQDSGGPLVCSIQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENG 401
Db 381 DSCQDSGGPLVCSIQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENG 440
QY 402 A 402
Db 441 A 441

RESULT 4

JN0560
u-plasminogen activator (EC 3.4.21.73) precursor - bovine
N:Alternate names: uPA
C:Species: Bos primigenius taurus (cattle)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:Accession: JN0560
R:Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A:Title: Bovine urokinase-type plasminogen activator and its receptor: cloning
A:Reference number: JN0560; MUID:93216119; PMID:8385052
A:Accession: JN0560

A:Accession: I52209
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 145-161 <NAG1>
A:Cross-references: GB:K03027; NID:G340174; PIDN:AAA61257.1; PID:G340175
R:Nagai, M.; Hiratsuku, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama, Gene 36, 183-188, 1985
A:Title: Molecular cloning of cDNA coding for human preprourokinase.
A:Reference number: J0102; MUID:86056954; PMID:2415429
A:Accession: J0102
A:Molecule type: mRNA
A:Residues: 1-213, 'I', 215-431 <NAG2>
A:Cross-references: GB:K03026; NID:G340155; PIDN:AA097138.1; PID:G340158; GB:D00244; NID:G340175
R:Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A:Title: Identification and primary sequence of an unspliced human urokinase poly(A) + RN
A:Reference number: A37561; MUID:84272706; PMID:6589820
A:Accession: A37561
A:Molecule type: mRNA
A:Residues: 66-431 <VER>
A:Cross-references: GB:D00244; NID:G220138
R:Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elsen, DNA 4, 139-146, 1985
A:Title: Molecular cloning, sequencing, and expression in *Escherichia coli* of human pre
A:Reference number: I38102; MUID:85203359; PMID:3888571
A:Accession: I38102
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>
A:Cross-references: EMBL:X02760; NID:G35297; PIDN:CAA56535.1; PID:G35298
R:Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996
A:Title: Characterization of single chain urokinase-type plasminogen activator with a no
A:Reference number: S65783; MUID:96186279; PMID:8652631
A:Accession: S65783
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>
A:Cross-references: EMBL:D11143; NID:G311467; PIDN:BA01919.1; PID:G119928
R:Günzler, W.A.; Steffens, G.J.; Otting, F.; Klm, S.M.A.; Frankus, E.; Flohe, L.
Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
A:Title: The primary structure of high molecular mass urokinase from human urine.
A:Reference number: A37562; MUID:83055084; PMID:6754569
A:Accession: A37562
A:Molecule type: protein
A:Residues: 21-177 <GUN>
R:Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O.
Eur. J. Biochem. 125, 251-257, 1982
A:Title: Human low-molecular-weight urinary urokinase. Partial characterization and pred
A:Reference number: A37563; MUID:83003608; PMID:6749491
A:Accession: A37563
A:Molecule type: protein
A:Residues: 156-176; 179-193, 'T', 195, 'T', 197-224 <SCH>
R:Steffens, G.J.; Günzler, W.A.; Otting, F.; Frankus, E.; Flohe, L.
Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
A:Title: The complete amino acid sequence of low molecular mass urokinase from human ur
A:Reference number: A37564; MUID:83055099; PMID:6754572
A:Accession: A37564
A:Molecule type: protein
A:Residues: 158-410 <STE>
R:Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K.
Biochem. Biophys. Res. Commun. 171, 401-406, 1990
A:Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinan
A:Reference number: A35689; MUID:90365737; PMID:2393398
A:Accession: A35689
A:Molecule type: protein
A:Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>
A:Note: Identification of a fucose and attempt to determine its attachment site
R:Rabani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltz Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
A:Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li
A:Reference number: A36697; MUID:91097529; PMID:2125213
A:Accession: A36697

A:Molecule type: protein
A:Residues: 21-34 <RAB>
R:Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.
submitted to the Brookhaven Protein Data Bank, July 1993
A:Reference number: A51255; PDB:1KDU
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
R:Li, X.; Smith, R.A.G.; Dobson, C.M.
Biochemistry 31, 9562-9571, 1992
A:Title: Sequential (1)H NMR assignments and secondary structure of the kringle
A:Reference number: A44375; MUID:93003110; PMID:1327118
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
R:Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazar, A.P.; Ole
submitted to the Brookhaven Protein Data Bank, January 1994
A:Reference number: A66922; PDB:1URK
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
R:Spragon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobso
submitted to the Brookhaven Protein Data Bank, July 1995
A:Reference number: A66058; PDB:1LMW
A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175;
C:Comment: This enzyme is found in urine in a high molecular mass form, consist
C:Comment: Urokinase-type plasminogen activator proteolytically activates plasmin
C:Genetics:
A:Gene: GDB:PLAU
A:Cross-references: GDB:119497; OMIM:191840
A:Map position: 10q24-10q24
A:Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
C:Function:
A:Description: proteolytically activates plasminogen
A:Pathway: fibrinolysis
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homo
C:Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-431/Product: urokinase-type plasminogen activator, single chain form #stat
F:21-177/Product: urokinase-type plasminogen activator chain A #status experime
F:31-62/Domain: EGF homology <EGF>
F:70-151/Domain: kringle homology <KRG>
F:156-177/Product: urokinase-type plasminogen activator chain A #status experi
F:179-431/Product: urokinase-type plasminogen activator chain B #status experi
F:179-419/Domain: tryptase homology <TRY>
F:31-39, 51, 53-62, 70-151, 91-133, 122-146, 168-299, 209-225, 217-288, 313-382, 345-3
F:38/Binding site: carbohydrate (Thr) (covalent) #status predicted
F:178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
F:224, 275, 376/Active site: His, Asp, Ser #status experimental
F:322/Binding site: carbohydrate (Asn) (covalent) #status experimental
Query Match 99.2%; Score 2240; DB 1; Length 431;
Best Local Similarity 97.8%; Pred. No. 2, 4e-166;
Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIDKSKTCYEGNGHFFRG 60
Db 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIDKSKTCYEGNGHFFRG 80
QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRPPWCYVQVGLK 120
Db 81 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRPPWCYVQVGLK 140
QY 121 PLVQECMWHDCADGX-----LKFQCGQKTLRPRFKIIGGEFTTIENQPFALIRRH 172
Db 141 PLVQECMWHDCADGXKPPSPBELXFCQCGKTLRPRFKIIGGEFTTIENQPFALIRRH 200
QY 173 RGSVTVYCGGSLSPCWVISATHCFIDYPKKEDIVYLGRLNSNTQGMKFEVENLI 232
Db 201 RGSVTVYCGGSLSPCWVISATHCFIDYPKKEDIVYLGRLNSNTQGMKFEVENLI 260
QY 233 LHMYSADTLAHNDIALKIRSKGRCAQPSRTITICLPNMYNDPQGTSCITGFGK 292
Db 261 LHMYSADTLAHNDIALKIRSKGRCAQPSRTITICLPNMYNDPQGTSCITGFGK 320
QY 293 ENSTDLYPEQLKMTVVKLISHRECOOPHYGVSEVTTKMLCAADPOKWTDSQCGSGGFL 352
Db 321 ENSTDLYPEQLKMTVVKLISHRECOOPHYGVSEVTTKMLCAADPOKWTDSQCGSGGFL 380

d) - human
 ase; urokinase-type I
 urokinase-type plasmin
 geinogen activator
 change 15-Sep-2000
 3; A37563; A37563; A
 , S.; Blasi, F.
 d its promoter.

417 SRCQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRC 476
191 QOPHYGSEVTTKMLCAAD-----POMKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRG 244
477 TSQHLNRTVTYDNNMLCAGDTRSGGPGQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLG 536
245 CALKDKPGVYTRVSHFLPWIRSHTK 269
537 CGQKDVPGVYTKVTNYLDWIRDNR 561

Search completed: May 25, 2004, 15:00:03
Job time : 21.1096 secs

Query Match 38.7%; Score 583; DB 6; Length 562;
Best Local Similarity 44.9%; Pred. No. 6.1e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO:3:
; LENGTH: 562
5185259-3

13 CG-QKTLRPRFKIIGGFTTIENQWPFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHC 70
299 CGLRQYSQOPFRKGGFLADIASHPWQAIFAKHRRSPGERFLCGGILISSCWILSAHC 358
71 FIDYPKKEDYIVLGRSLNSNTQGENKFEVENILHKYSDATLAHNDIALLKIRSK 130
359 FQERFPPHLLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDS 416
131 GRCAQPSRTIOTICLPSMYNDPQFGTSCITGFGKENSTDYLYPEOLKMTVVKLIHSHREC 190
417 SRCQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRC 476
191 QOPHYGSEVTTKMLCAAD-----POMKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRG 244
477 TSQHLNRTVTYDNNMLCAGDTRSGGPGQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLG 536
245 CALKDKPGVYTRVSHFLPWIRSHTK 269
537 CGQKDVPGVYTKVTNYLDWIRDNR 561

RESULT 50
5200340-2
; Patent No. 5200340
; APPLICANT: POSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,
; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
; ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
; SEQ ID NO:2:
; LENGTH: 562
5200340-2

Query Match 38.7%; Score 583; DB 6; Length 562;
Best Local Similarity 44.9%; Pred. No. 6.1e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO:3:
; LENGTH: 562
5185259-3

13 CG-QKTLRPRFKIIGGFTTIENQWPFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHC 70
299 CGLRQYSQOPFRKGGFLADIASHPWQAIFAKHRRSPGERFLCGGILISSCWILSAHC 358
71 FIDYPKKEDYIVLGRSLNSNTQGENKFEVENILHKYSDATLAHNDIALLKIRSK 130
359 FQERFPPHLLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDS 416
131 GRCAQPSRTIOTICLPSMYNDPQFGTSCITGFGKENSTDYLYPEOLKMTVVKLIHSHREC 190

417 SRCQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRC 476
191 QOPHYGSEVTTKMLCAAD-----POMKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRG 244
477 TSQHLNRTVTYDNNMLCAGDTRSGGPGQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLG 536
245 CALKDKPGVYTRVSHFLPWIRSHTK 269
537 CGQKDVPGVYTKVTNYLDWIRDNR 561

Search completed: May 25, 2004, 15:00:03
Job time : 21.1096 secs

Query Match 38.7%; Score 583; DB 6; Length 562;
Best Local Similarity 44.9%; Pred. No. 6.1e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO:3:
; LENGTH: 562
5185259-3

13 CG-QKTLRPRFKIIGGFTTIENQWPFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHC 70
299 CGLRQYSQOPFRKGGFLADIASHPWQAIFAKHRRSPGERFLCGGILISSCWILSAHC 358
71 FIDYPKKEDYIVLGRSLNSNTQGENKFEVENILHKYSDATLAHNDIALLKIRSK 130
359 FQERFPPHLLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDS 416
131 GRCAQPSRTIOTICLPSMYNDPQFGTSCITGFGKENSTDYLYPEOLKMTVVKLIHSHREC 190

Best Local Similarity 44.9%; Pred. No. 6.1e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
QY 13 CG-QKTLRPRFKIIGGEFTTIENQPFALYRRH-RGGSVTYVCGSLISPCWISATHC 70
Db 299 CGLRQYQOPFRKIGGLFADIAHPWQAALFAKRRSPGERFLCGGILISSCWILSAHC 358
QY 71 FIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSK 130
Db 359 FQERPPPHLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDS 416
QY 131 GRCAQPSRTIOTICLPSMYNDPOFGTSCEITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
Db 417 SRCAQESSVVRTVCLPPADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRC 476
QY 191 QOPHYGSEVTTKMLCAAD-----POWKT-DSQCGSDGGPLVCSLQGRMTLTGIVSWG 244
Db 477 TSQHLLNRTVTDNMLCAGDTRSGGPQANLHDACQSDGGPLVCLNDGRMTLVGIISWGLG 536
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
Db 537 CGQKDVPGVYTKVTNYLDWIRDNR 561
RESULT 48
US-09-703-695A-4
; Sequence 4, Application US/09703695A
; Patent No. 6593097
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: P1788R1
; CURRENT APPLICATION NUMBER: US/09/703,695A
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-703-695A-4
Query Match 38.7%; Score 583; DB 4; Length 562;
Best Local Similarity 44.9%; Pred. No. 6.1e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
QY 13 CG-QKTLRPRFKIIGGEFTTIENQPFALYRRH-RGGSVTYVCGSLISPCWISATHC 70
Db 299 CGLRQYQOPFRKIGGLFADIAHPWQAALFAKRRSPGERFLCGGILISSCWILSAHC 358
QY 71 FIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSK 130
Db 359 FQERPPPHLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDS 416
QY 131 GRCAQPSRTIOTICLPSMYNDPOFGTSCEITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
Db 417 SRCAQESSVVRTVCLPPADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRC 476
QY 191 QOPHYGSEVTTKMLCAAD-----POWKT-DSQCGSDGGPLVCSLQGRMTLTGIVSWG 244
Db 477 TSQHLLNRTVTDNMLCAGDTRSGGPQANLHDACQSDGGPLVCLNDGRMTLVGIISWGLG 536
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
Db 537 CGQKDVPGVYTKVTNYLDWIRDNR 561
RESULT 49
5185259-3
; Patent No. 5185259
; APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHAR, GORDON A.

STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-50
Query Match 38.7%; Score 583; DB 2; Length 562;
Best Local Similarity 44.9%; Pred. No. 6.1e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
QY 13 CG-QKTLRPRFKIIGGEFTTIENQPFALYRRH-RGGSVTYVCGSLISPCWISATHC 70
Db 299 CGLRQYQOPFRKIGGLFADIAHPWQAALFAKRRSPGERFLCGGILISSCWILSAHC 358
QY 71 FIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSK 130
Db 359 FQERPPPHLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDS 416
QY 131 GRCAQPSRTIOTICLPSMYNDPOFGTSCEITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
Db 417 SRCAQESSVVRTVCLPPADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRC 476
QY 191 QOPHYGSEVTTKMLCAAD-----POWKT-DSQCGSDGGPLVCSLQGRMTLTGIVSWG 244
Db 477 TSQHLLNRTVTDNMLCAGDTRSGGPQANLHDACQSDGGPLVCLNDGRMTLVGIISWGLG 536
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
Db 537 CGQKDVPGVYTKVTNYLDWIRDNR 561
RESULT 47
US-08-883-795A-38
; Sequence 38, Application US/08883795A
; Patent No. 5985607
; GENERAL INFORMATION:
; APPLICANT: Delcuve, Genevieve
; APPLICANT: Awang, Gregor
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/883,795A
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Michelle
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 7841-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 562 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein (rpa)
US-08-883-795A-38
Query Match 38.7%; Score 583; DB 2; Length 562;

```
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO: 8
; LENGTH: 527
; 5185259-8

Query Match      38.7%; Score 583; DB 6; Length 527;
Best Local Similarity 44.9%; Pred. No. 5.6e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPFKIIIGGFTTIENQWFAAIYRRH-RGGSVTVYVCGSLISPCWVISATHC 70
Db 264 CGLRQYSQPFRIKGGFLADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 323
QY 71 FIDYPPKEDYIVYIGRSRLNSNTQGEKMEVENLILHKDYSADTLAHNDIALKIRSK 130
Db 324 FQERFPPHLLTVILGRTYRVVPGEEQKFEVEKIYVHKFDDDT--YDNDIALQLKSDS 381
QY 131 GRCAQPSRTITQICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLIHSHREC 190
Db 382 SRCQESSVVRTVCLPPADQLPDWTECELSGYGKHEALSPFYSERLKEAHRVLYPSSRC 441
QY 191 QQPHYGSEVTTKMLCAAD-----PQWKT-DSQCGSDGGLVCSLQGRMTLTGIVSWG 244
Db 442 TSQHLNRTVTDNMLCAGDTRSGGPOANLHDACQSDSGGLVCLNDGRMTLVGIISWGLG 501
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
Db 502 CGQKDPGVYTKVTNYLDWIRDNR 526

RESULT 45
US-08-811-949-43
; Sequence 43, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORVAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220

; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 562 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-43

Query Match      38.7%; Score 583; DB 2; Length 562;
Best Local Similarity 44.9%; Pred. No. 6.1e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPFKIIIGGFTTIENQWFAAIYRRH-RGGSVTVYVCGSLISPCWVISATHC 70
Db 299 CGLRQYSQPFRIKGGFLADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 358
QY 71 FIDYPPKEDYIVYIGRSRLNSNTQGEKMEVENLILHKDYSADTLAHNDIALKIRSK 130
Db 359 FQERFPPHLLTVILGRTYRVVPGEEQKFEVEKIYVHKFDDDT--YDNDIALQLKSDS 416
QY 131 GRCAQPSRTITQICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLIHSHREC 190
Db 417 SRCQESSVVRTVCLPPADQLPDWTECELSGYGKHEALSPFYSERLKEAHRVLYPSSRC 476
QY 191 QQPHYGSEVTTKMLCAAD-----PQWKT-DSQCGSDGGLVCSLQGRMTLTGIVSWG 244
Db 477 TSQHLNRTVTDNMLCAGDTRSGGPOANLHDACQSDSGGLVCLNDGRMTLVGIISWGLG 536
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
Db 537 CGQKDPGVYTKVTNYLDWIRDNR 561

RESULT 46
US-08-560-098A-50
; Sequence 50, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEB, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 562 amino acids
; TYPE: amino acid
```

Db 387 TSQHLLNRTVTDNMLCAGDTRSGGPOANLHDACQSDGGLVCLNDGRMTLVGIISWGLG 446
QY 245 CALKDKPGVYTVRSHFLPWIRSHTK 269
Db 447 CGQKDPGVYTVKVTNYLDWIRDNR 471

RESULT 42

US-07-609-510B-16
; Sequence 16, Application US/07609510B
; Patent No. 5326700
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue PI
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN.
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,510B
; FILING DATE: 19901106
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 38.7%; Score 583; DB 1; Length 527;

Best Local Similarity 44.9%; Pred. No. 5.6e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLPRFKIIGGEFTTIENQWFAAIYRRH-RGGSVTVVCGGSLISPCWVISATHC 70
Db 264 CGLRQYSQPFRIKGLFADIAHPWQAAIFAKHRSFGERFLCGGILISSCWILSAHC 323
QY 71 FIDYPKKEDYIYVLRSLNNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
Db 324 FQERPPPHLTIVLGRYRVVPGEBEQFEVEKYIVHKEFDDT--YNDNDIALQLKSDS 381
QY 131 GRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEOLKMTVVKLISHREC 190
Db 382 SRCQESSVVRVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSRC 441
QY 191 QCPHYYGSEVTTMLCAAD-----PQWKT-DSCQDGGPLVCSLQGRMTLTGIVSWG 244
Db 442 TSQHLNRTVTDNMLCAGDTRSGGPOANLHDACQSDGGLVCLNDGRMTLVGIISWGLG 501
QY 245 CALKDKPGVYTVRSHFLPWIRSHTK 269
Db 502 CGQKDPGVYTVKVTNYLDWIRDNR 526

RESULT 43

PCT-US91-01025A-2
; Sequence 2, Application PC/TUS9101025A
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
; TITLE OF INVENTION: Specific Properties

; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01025A
; FILING DATE: 19910214
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/486,657
; FILING DATE: 1 March 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 454P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1896
; TELEFAX: 415/352-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PCT-US91-01025A-2
Query Match 38.7%; Score 583; DB 5; Length 527;
Best Local Similarity 44.9%; Pred. No. 5.6e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
QY 13 CG-QKTLPRFKIIGGEFTTIENQWFAAIYRRH-RGGSVTVVCGGSLISPCWVISATHC 70
Db 264 CGLRQYSQPFRIKGLFADIAHPWQAAIFAKHRSFGERFLCGGILISSCWILSAHC 323
QY 71 FIDYPKKEDYIYVLRSLNNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
Db 324 FQERPPPHLTIVLGRYRVVPGEBEQFEVEKYIVHKEFDDT--YNDNDIALQLKSDS 381
QY 131 GRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEOLKMTVVKLISHREC 190
Db 382 SRCQESSVVRVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSRC 441
QY 191 QCPHYYGSEVTTMLCAAD-----PQWKT-DSCQDGGPLVCSLQGRMTLTGIVSWG 244
Db 442 TSQHLNRTVTDNMLCAGDTRSGGPOANLHDACQSDGGLVCLNDGRMTLVGIISWGLG 501
QY 245 CALKDKPGVYTVRSHFLPWIRSHTK 269
Db 502 CGQKDPGVYTVKVTNYLDWIRDNR 526
RESULT 44
5185259-8
; Patent No. 5185259
; APPLICANT: GOEDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHAR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694

RESULT 40

; Sequence 6, Application US/09410882

; GENERAL INFORMATION

; TITLE OF INVENTION

; CORRESPONDENCE ADDRESS:

STREET: 60 State Street

STATE: MA

ZIP: 02109
COMPTON 223 02109

MEDIUM TYPE: TOLBY &
COMPUTER: IBM PC COME

```

TEACHING SYSTEM:  FC
SOFTWARE:  ASCIT (text

```

APPLICATION NUMBER: U

CLASSIFICATION:

APPLICATION NUMBER: U

APPLICATION NUMBER: U

ATTORNEY/AGENT INFORMATION

REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER:

TELEPHONE: (617) 227-

; INFORMATION FOR SEQ ID NO:

; LENGTH: 383 amino acids

; TOPOLOGY: linear

US-09-410-882-6

Query Match

Matchless 115; Combezvalive

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

..

[illegible]

Patent No. 6309861
GENERAL INFORMATION:
APPLICANT: Ambrosius, Dorothee
APPLICANT: Rudolph, Rainer
APPLICANT: Schaeffner, Joerg
APPLICANT: Schwarz, Elisabeth
TITLE OF INVENTION: Processes for the production of naturally folded and secreted proteins
FILE REFERENCE: Case 20379
CURRENT APPLICATION NUMBER: US/09/553,498
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: EP99107412.1
PRIOR FILING DATE: 1999-04-26
NUMBER OF SEQ ID NOS: 10
SEQ ID NO 10
LENGTH: 378
TYPE: PRT
ORGANISM: E. coli
US-09-553-498-10

Query Match 38.7%; Score 583; DB 4; Length 378;
Best Local Similarity 44.9%; Pred. No. 3.5e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGGEFTTIENOPWFAAIYRRH-RGGSVTYVCGGSLISPCWVISATHC 70
DB 115 CGLRQYSQOPRIKGLFADIASHPWQAAIFAKHRSRSPGERFLCGGILISSCWILSAHC 174
QY 71 FIDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130
DB 175 FQERPPPHLTVILGRYRVVPGGEEQKFEVEKIVHKEFDDT--YNDIALIQLKSDS 232
QY 131 GRCAQPSRTIOTICLPMSYNDPQFCTSCETGFGKENSVDLYPEQLKMTVVKLISHREC 190
DB 233 SRCQESSVVRIVCLPPADLQLPDWTCELSYKHEALSPFYSERLKEAHVRLYPSRC 292
QY 191 QOPHYGSEVTTKMLCAAD-----PWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRG 244
DB 293 TSQHLLNRTVDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 352
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB 353 CGQKDVPGVYTKVNYLDWIRDNR 377

RESULT 38
US-09-618-869-10
Sequence 10, Application US/09618869
Patent No. 6455279
GENERAL INFORMATION:
APPLICANT: Ambrosius, Dorothee
APPLICANT: Rudolph, Rainer
APPLICANT: Schaeffner, Joerg
APPLICANT: Schwarz, Elisabeth
TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR CHAPERONES
FILE REFERENCE: 20381
CURRENT APPLICATION NUMBER: US/09/618,869
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: EP99114811.5
PRIOR FILING DATE: 1999-07-29
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 378
TYPE: PRT
ORGANISM: Escherichia coli
US-09-618-869-10

Query Match 38.7%; Score 583; DB 4; Length 378;
Best Local Similarity 44.9%; Pred. No. 3.5e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGGEFTTIENOPWFAAIYRRH-RGGSVTYVCGGSLISPCWVISATHC 70
DB 115 CGLRQYSQOPRIKGLFADIASHPWQAAIFAKHRSRSPGERFLCGGILISSCWILSAHC 174
QY 71 FIDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130
DB 175 FQERPPPHLTVILGRYRVVPGGEEQKFEVEKIVHKEFDDT--YNDIALIQLKSDS 232
QY 131 GRCAQPSRTIOTICLPMSYNDPQFCTSCETGFGKENSVDLYPEQLKMTVVKLISHREC 190
DB 233 SRCQESSVVRIVCLPPADLQLPDWTCELSYKHEALSPFYSERLKEAHVRLYPSRC 292
QY 191 QOPHYGSEVTTKMLCAAD-----PWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRG 244
DB 293 TSQHLLNRTVDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 352
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB 353 CGQKDVPGVYTKVNYLDWIRDNR 377

RESULT 39
US-08-558-269-6
Sequence 6, Application US/08558269
Patent No. 5961973
GENERAL INFORMATION:
APPLICANT: Crea, Roberto
TITLE OF INVENTION: PATHOGEN-TARGETED BIOCATALYSTS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/558,269
FILING DATE: 13-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,800
FILING DATE: 08-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CRI-001CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 383 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-558-269-6

Query Match 38.7%; Score 583; DB 2; Length 383;
Best Local Similarity 44.9%; Pred. No. 3.6e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGGEFTTIENOPWFAAIYRRH-RGGSVTYVCGGSLISPCWVISATHC 70
DB 120 CGLRQYSQOPRIKGLFADIASHPWQAAIFAKHRSRSPGERFLCGGILISSCWILSAHC 179
QY 71 FIDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130

; Sequence 10, Application US/09553498


```

; ; NUMBER OF SEQUENCES: 3
; ; CURRENT APPLICATION DATA:
; ; APPLICATION NUMBER: US/07/585,129
; ; FILING DATE: 06-FEB-1990
; ; SEQ ID NO:1:
; ; LENGTH: 355
5223256-1

Query Match 38.7%; Score 583; DB 6; Length 355;
Best Local Similarity 44.9%; Pred. No. 3,2e-52;

```

Db 270 TSQHLLNRTVTDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 329

QY 245 CALXDKPGVYTRVSHFLPWIRSHTK 269

Db 330 CGQKDVPGVYTKVTNYLDWIRDNR 354

RESULT 30

US-08-217-617A-1

; Sequence 1, Application US/08217617A

; Patent No. 5676947

; GENERAL INFORMATION:

; APPLICANT: Martin, Ulrich

; APPLICANT: Fischer, Stephan

; TITLE OF INVENTION: METHOD FOR TREATING THROMBOEMBOLIC

; TITLE OF INVENTION: CONDITIONS USING THROMBOLYTICALLY ACTIVE PROTEINS

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch

; STREET: 805 Third Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/217,617A

; FILING DATE: 25-MARCH-1994

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 5676947man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: BOER 1030.1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 355 amino acids

; TYPE: amino acids

; TOPOLOGY: linear

US-08-217-617A-1

Query Match 38.7%; Score 583; DB 1; Length 355;

Best Local Similarity 44.9%; Pred. No. 3.2e-52;

Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-OKTLRPRFKIIGGFTTIENOPWFAAIYRRH-RGGSVTVYVCGGLISPCWVISATHC 70

Db 92 CGLRQYQOPQFRIGKGLFADIAHPWQAIFAKHRRSPGERFLCGGILISSCWILSAHC 151

QY 71 FIDYPKKEDYIVYLGSRSLNSNTQGMKFVENILHKOYSADTLAHNDIALKIRSK 130

Db 152 FQERPPPHLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDS 209

QY 131 GRCAQPSRTIOTICLPMSYNDPQGTSCETITGFGKNSDLYLPEQLKMTVVKLISHREC 190

Db 210 SRCAQESSVVRTVCLPPADLQLPDWTCELSGKGHEALSPFYSERLKEAHVLYPSSRC 269

QY 191 QQPHYGSEVTKMLCAAD-----PWKT-DSOQDGGPLVCSLQGRMTLTGIVSWGRG 244

Db 270 TSQHLLNRTVTDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 329

QY 245 CALXDKPGVYTRVSHFLPWIRSHTK 269

Db 330 CGQKDVPGVYTKVTNYLDWIRDNR 354

RESULT 31

US-08-811-949-45

; Sequence 45, Application US/08811949

; Patent No. 5840533

; GENERAL INFORMATION:

; APPLICANT: NIWA, MINEO

; APPLICANT: SAITO, YOSHIMASA

US-08-217-616-1

; Sequence 1, Application US/08217616

; Patent No. 5690931

; GENERAL INFORMATION:

; APPLICANT: Martin, Ulrich

; APPLICANT: Fischer, Stephan

; TITLE OF INVENTION: METHOD FOR TREATING THROMBOEMBOLIC CONDITIONS

; TITLE OF INVENTION: BY INHIBITING REOCCLUSION VIA THE USE OF

; TITLE OF INVENTION: MULTIPLE BOLUS ADMINISTRATION OF

; TITLE OF INVENTION: THROMBOLYTICALLY ACTIVE PROTEINS

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch

; STREET: 805 Third Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/217,616

; FILING DATE: 25-MARCH-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 5690931man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: BOER 1026.1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 355 amino acids

; TYPE: amino acids

; TOPOLOGY: linear

US-08-217-616-1

Query Match 38.7%; Score 583; DB 1; Length 355;

Best Local Similarity 44.9%; Pred. No. 3.2e-52;

Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5

QY 13 CG-OKTLRPRFKIIGGFTTIENOPWFAAIYRRH-RGGSVTVYVCGGLISPCWVISATHC 70

Db 92 CGLRQYQOPQFRIGKGLFADIAHPWQAIFAKHRRSPGERFLCGGILISSCWILSAHC 151

QY 71 FIDYPKKEDYIVYLGSRSLNSNTQGMKFVENILHKOYSADTLAHNDIALKIRSK 130

Db 152 FQERPPPHLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDS 209

QY 131 GRCAQPSRTIOTICLPMSYNDPQGTSCETITGFGKNSDLYLPEQLKMTVVKLISHREC 190

Db 210 SRCAQESSVVRTVCLPPADLQLPDWTCELSGKGHEALSPFYSERLKEAHVLYPSSRC 269

QY 191 QQPHYGSEVTKMLCAAD-----PWKT-DSOQDGGPLVCSLQGRMTLTGIVSWGRG 244

Db 270 TSQHLLNRTVTDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 329

QY 245 CALXDKPGVYTRVSHFLPWIRSHTK 269

Db 330 CGQKDVPGVYTKVTNYLDWIRDNR 354

RESULT 32

US-08-811-949-45

; Sequence 45, Application US/08811949

; Patent No. 5840533

; GENERAL INFORMATION:

; APPLICANT: NIWA, MINEO

; APPLICANT: SAITO, YOSHIMASA

```

; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acids
; TOPOLOGY: linear
US-08-217-618-1

Query Match 38.7%; Score 583; DB 1; Length 355;
Best Local Similarity 44.9%; Pred. No. 3.2e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRKIIIGGEFTTIENQWFAAIVRRH-RGGSVTVVCGGSLISPCWVISATHC 70
Db 92 CGLRQYSQPQRIKGGFLADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 151
QY 71 FIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
Db 152 FOERPPPHLTIVILGRTRVVPGEEOQFEVEKIVHKEFDDDT--YNDIALQLKSDS 209
QY 131 GRCAQPSRTIOTICLPSMYNDPQFCTSCIEITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
Db 210 SRCQESSVVRTVCLPPADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRC 269
QY 191 QOPHYGSEVTTKMLCAAD-----PWKT-DSCQDGGPLVCSLQGRMTLTGIVSWG 244
Db 270 TSOHLNRTVTDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 329
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
Db 330 CGQKDVPGVYTKVNYLDWIRDNR 354

RESULT 28
US-08-427-640-2
; Sequence 2, Application US/08427640
; Patent No. 5658788
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Compounds and Methods for Treatment of Thromboembolic Disor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN.
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,640
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/689,410
; FILING DATE: 22 APRIL 1991
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-427-640-2

Query Match 38.7%; Score 583; DB 1; Length 355;
Best Local Similarity 44.9%; Pred. No. 3.2e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRKIIIGGEFTTIENQWFAAIVRRH-RGGSVTVVCGGSLISPCWVISATHC 70
Db 92 CGLRQYSQPQRIKGGFLADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 151
QY 71 FIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
Db 152 FOERPPPHLTIVILGRTRVVPGEEOQFEVEKIVHKEFDDDT--YNDIALQLKSDS 209
QY 131 GRCAQPSRTIOTICLPSMYNDPQFCTSCIEITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
Db 210 SRCQESSVVRTVCLPPADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRC 269
QY 191 QOPHYGSEVTTKMLCAAD-----PWKT-DSCQDGGPLVCSLQGRMTLTGIVSWG 244

```

```

Db 92 CGLRQYSQPQRIKGGFLADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 151
QY 71 FIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
Db 152 FOERPPPHLTIVILGRTRVVPGEEOQFEVEKIVHKEFDDDT--YNDIALQLKSDS 209
QY 131 GRCAQPSRTIOTICLPSMYNDPQFCTSCIEITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
Db 210 SRCQESSVVRTVCLPPADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRC 269
QY 191 QOPHYGSEVTTKMLCAAD-----PWKT-DSCQDGGPLVCSLQGRMTLTGIVSWG 244
Db 270 TSOHLNRTVTDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 329
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
Db 330 CGQKDVPGVYTKVNYLDWIRDNR 354

RESULT 29
US-08-427-640-6
; Sequence 6, Application US/08427640
; Patent No. 5658788
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Compounds and Methods for Treatment of Thromboembolic isor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN.
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,640
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/689,410
; FILING DATE: 22 APRIL 1991
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-427-640-6

Query Match 38.7%; Score 583; DB 1; Length 355;
Best Local Similarity 44.9%; Pred. No. 3.2e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRKIIIGGEFTTIENQWFAAIVRRH-RGGSVTVVCGGSLISPCWVISATHC 70
Db 92 CGLRQYSQPQRIKGGFLADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 151
QY 71 FIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
Db 152 FOERPPPHLTIVILGRTRVVPGEEOQFEVEKIVHKEFDDDT--YNDIALQLKSDS 209
QY 131 GRCAQPSRTIOTICLPSMYNDPQFCTSCIEITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
Db 210 SRCQESSVVRTVCLPPADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRC 269
QY 191 QOPHYGSEVTTKMLCAAD-----PWKT-DSCQDGGPLVCSLQGRMTLTGIVSWG 244

```

REFERENCE/DOCKET NUMBER: 18-966-0
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-57

Query Match 38.9%; Score 587; DB 2; Length 437;
Best Local Similarity 45.3%; Pred. No. 1.6e-52;
Matches 120; Conservative 38; Mismatches 97; Indels 10; Gaps 5;
QY 13 CG-QKTLRPFKIIIGGFTTIENQWFAAIYRRH-RGGSVTVYVCGSLISPCWVISATHC 70
DB 174 CGLRQVSGQDFDIIGLFDIASHPWQAIFAKHRRSPGERFLCGGILISSCWILSAHC 233
QY 71 FIDYPKEDYIVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
DB 234 FQERFPPHLLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDS 291
QY 131 GRCAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEOLKMTVVKLIHREC 190
DB 292 SRCQAESVVRVCLPADLQLPDWTCELSGKGHEALSPFYSERLKEAHLVPSRRC 351
QY 191 QPHYGVSEVTTKMLCAAD-----PQWKT-DSQGDGGPLVCSLQGRMTLGTIVSWG 244
DB 352 TSQHLNRTVTDNMLCAGDTRSGGPQANLHDAQCQDGGPLVCLNDGRMTLGTIVSWG 411
QY 245 CALKDKGVYTRVSHFLPWIRSHTK 269
DB 412 CGQKDPGVYTKVNTYLDWIRDMR 436

RESULT 26

US-08-137-116-1
Sequence 1, Application US/08137116

Patent No. 5500411
GENERAL INFORMATION:
APPLICANT: Martin, Ulrich
TITLE OF INVENTION: METHOD FOR TREATING THROMBOEMBOLIC
CONDITIONS BY INHIBITING REOCCCLUSION
TITLE OF INVENTION: CONDITIONS BY INHIBITING REOCCCLUSION
TITLE OF INVENTION: VIA THE USE OF MULTIPLE BODIES
TITLE OF INVENTION: ADMINISTRATION OF THROMBOLYTICALLY
ACTIVE PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,116
FILING DATE: June 30, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP92/00851
FILING DATE: 15 April 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 41 12 398
FILING DATE: 16 April 1991
PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 41 23 845
FILING DATE: 18 July 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 550041man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: BOER 1026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acids
TOPOLOGY: linear
US-08-137-116-1

Query Match 38.7%; Score 583; DB 1; Length 355;
Best Local Similarity 44.9%; Pred. No. 3.2e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
QY 13 CG-QKTLRPFKIIIGGFTTIENQWFAAIYRRH-RGGSVTVYVCGSLISPCWVISATHC 70
DB 92 CGLRQVSGQDFDIIGLFDIASHPWQAIFAKHRRSPGERFLCGGILISSCWILSAHC 151
QY 71 FIDYPKEDYIVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
DB 152 FQERFPPHLLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDS 205
QY 131 GRCAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEOLKMTVVKLIHREC 190
DB 210 SRCQAESVVRVCLPADLQLPDWTCELSGKGHEALSPFYSERLKEAHLVPSRRC 265
QY 191 QPHYGVSEVTTKMLCAAD-----PQWKT-DSQGDGGPLVCSLQGRMTLGTIVSWG 244
DB 270 TSQHLNRTVTDNMLCAGDTRSGGPQANLHDAQCQDGGPLVCLNDGRMTLGTIVSWG 329
QY 245 CALKDKGVYTRVSHFLPWIRSHTK 269
DB 330 CGQKDPGVYTKVNTYLDWIRDMR 354

RESULT 27

US-08-217-618-1
Sequence 1, Application US/08217618

Patent No. 5510330
GENERAL INFORMATION:
APPLICANT: Martin, Ulrich
TITLE OF INVENTION: COMBINATIONS OF THROMBOLYTICALLY ACTIVE
PROTEINS AND NON-HEPARIN ANTICOAGULANTS, AND USES THEREOF
TITLE OF INVENTION: COMBINATIONS OF THROMBOLYTICALLY ACTIVE
PROTEINS AND NON-HEPARIN ANTICOAGULANTS, AND USES THEREOF
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/217,618
FILING DATE: 25-MARCH-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5510330man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: BOER 1032
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200

APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-51

Query Match 39.1%; Score 590; DB 2; Length 437;
Best Local Similarity 45.3%; Pred. No. 9e-53;
Matches 120; Conservative 39; Mismatches 96; Indels 10; Gaps 5;
QY 13 CG-QKTLPRFKIIGGEFTTIENQWPAAIYRRH-RGGSVTYVCGSLISPCWVISATHC 70
DB 174 CGLRYSQPFRIIGLGFADIASHPWQAIAFAKRRSPGERFLCGGILISSCWILSAHC 233
QY 71 FIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
DB 234 FQERPPPHLTVILGRTYRVVFGEEQKFEVEKIVHKEFDDT--YNDIALQLKSDS 291
QY 131 GRCAQPSRTIOTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
DB 292 SRCAQESSVVRIVCLPPADQLPDWTECELSGYKHEALSPFYERLKEAHRVLYPSRC 351
QY 191 QOPHYGSEVTTKMLCAAD-----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRG 244
DB 352 TSQHLNRTVTDNMLCAGTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 411
QY 245 CALKDKPGVYTVSHPLWIRSHTK 269
DB 412 CGQKDPGVYTKVNYLDWIRDNR 436

RESULT 24

US-08-811-949-39
Sequence 39, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-39

Query Match 39.1%; Score 590; DB 2; Length 527;
Best Local Similarity 45.3%; Pred. No. 1e-52;
Matches 120; Conservative 39; Mismatches 96; Indels 10; Gaps 5;
QY 13 CG-QKTLPRFKIIGGEFTTIENQWPAAIYRRH-RGGSVTYVCGSLISPCWVISATHC 70
DB 264 CGLRYSQPFRIIGLGFADIASHPWQAIAFAKRRSPGERFLCGGILISSCWILSAHC 323
QY 71 FIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
DB 324 FQERPPPHLTVILGRTYRVVFGEEQKFEVEKIVHKEFDDT--YNDIALQLKSDS 381
QY 131 GRCAQPSRTIOTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
DB 382 SRCAQESSVVRIVCLPPADQLPDWTECELSGYKHEALSPFYERLKEAHRVLYPSRC 441
QY 191 QOPHYGSEVTTKMLCAAD-----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRG 244
DB 442 TSQHLNRTVTDNMLCAGTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 501
QY 245 CALKDKPGVYTVSHPLWIRSHTK 269
DB 502 CGQKDPGVYTKVNYLDWIRDNR 526

RESULT 25

US-08-811-949-57
Sequence 57, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-47

Query Match          39.1%; Score 590; DB 2; Length 355;
Best Local Similarity 45.3%; Pred. No. 6e-53;
Matches 120; Conservative 39; Mismatches 96; Indels 10; Gaps 5

QY 13 CG-QKTLRPRFKIIGGEFTTIENQPWFAAIYRRH-RGGSVTVCVCGSLISPCWVISATHC 70
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 92 CGLRQVSQPFRIICGLFADIASHPQAIAFAKHRSRSPGERFLCGGILLISSCWILSAHC 151
QY 71 FDYPKKEDYIVVLGRSRLNSNTQGEMKFEVENLILHKDYSDATLAHNDDIALLKIRSK 130
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 152 FQRFPFHLLTVILGTTRYVPGEEOKEVEKYIVHKEFFDDT--YNDNIALQLKSDS 209
QY 131 GRCAQPSRTITCLPSMNDPFGTSCEITGFGENSTDYLYPEQLKMTVVVKLIASHREC 190
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 210 SRCAQSSSVRVTCVLPADLQPDWTECELSYGKHEALSPYSERLKEAHVRLYPSSRC 269
QY 191 QPHYYGVSEVTTKMLCAAD-----PWKT-DSOQSGGGPLVCSIQGRMTLTGIVSWG 244
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 270 TSQHLLNRVTDNMLCAGDTRSGGPQANLHDACQDGSGGPLVCLNDGRMTLVGIISWG 329
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 330 CGQKDVGVTYKTYNLWDIRDNR 354

RESULT 22
US-08-811-949-47
; Sequence 47, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; ADDRESS: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-59

Query Match          40.3%; Score 607; DB 2; Length 355;
Best Local Similarity 46.8%; Pred. No. 1e-54;
Matches 124; Conservative 36; Mismatches 95; Indels 10; Gaps 5

QY 13 CG-QKTLRPRFKIIGGEFTTIENQPWFAAIYRRH-RGGSVTVCVCGSLISPCWVISATHC 70
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 92 CGLRQVSQPFRIICGLFADIASHPQAIAFAKHRSRSPGERFLCGGILLISSCWILSAHC 151
QY 71 FDYPKKEDYIVVLGRSRLNSNTQGEMKFEVENLILHKDYSDATLAHNDDIALLKIRSK 130
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 152 FQRFPFHLLTVILGTTRYVPGEEOKEVEKYIVHKEFFDDT--YNDNIALQLKSDS 209
QY 131 GRCAQPSRTITCLPSMNDPFGTSCEITGFGENSTDYLYPEQLKMTVVVKLIASHREC 190
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 210 SRCAQSSSVRVTCVLPADLQPDWTECELSYGKHEALSPYSERLKEAHVRLYPSSRC 269
QY 191 QPHYYGVSEVTTKMLCAAD-----PWKT-DSOQSGGGPLVCSIQGRMTLTGIVSWG 244
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 270 TSQHLLNRVTDNMLCAGDTRSGGPQANLHDACQDGSGGPLVCLNDGRMTLVGIISWG 329
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 330 CGQKDVGVTYKTYNLWDIRDNR 354

RESULT 22
US-08-811-949-47
; Sequence 47, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; ADDRESS: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-49

Query Match 91.1%; Score 1374; DB 2; Length 254;
Best Local Similarity 99.2%; Pred. No. 1.5e-134;
Matches 252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 23 KIIGFTTIENQPFPAIYRRHGGSVTVVCGSLISPCWVISATHCFIDYPKKEDIVY 82
DB 1 KIIGFTTIENQPFPAIYRRHGGSVTVVCGSLISPCWVISATHCFIDYPKKEDIVY 60
QY 83 YLGRSLNNTQGMKFEVENILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQT 142
DB 61 YLGRSLNNTQGMKFEVENILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQT 120
QY 143 ICLPSMYNDPQFTSCETIGFGKENSTDYLYPEQLKMTVVKLIISHRECOQPHYGVSEVTT 202
DB 121 ICLPSMYNDPQFTSCETIGFGKENSTDYLYPEQLKMTVVKLIISHRECOQPHYGVSEVTT 180
QY 203 KMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHFLP 262
DB 181 KMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHFLP 240
QY 263 WIRSHTKEENGLAL 276
DB 241 WIRSHTKEENGLVL 254

RESULT 21
US-08-811-949-59
; Sequence 59, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SAKAI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
; US-08-944-483-73

Query Match 91.6%; Score 1382; DB 3; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.2e-135;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 LIIGFTTIENQPFPAIYRRHGGSVTVVCGSLISPCWVISATHCFIDYPKKEDIVY 83
DB 1 LIIGFTTIENQPFPAIYRRHGGSVTVVCGSLISPCWVISATHCFIDYPKKEDIVY 60
QY 84 LGRSLNNTQGMKFEVENILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQT 143
DB 61 LGRSLNNTQGMKFEVENILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQT 120
QY 144 CLPSMYNDPQFTSCETIGFGKENSTDYLYPEQLKMTVVKLIISHRECOQPHYGVSEVTTK 203
DB 121 CLPSMYNDPQFTSCETIGFGKENSTDYLYPEQLKMTVVKLIISHRECOQPHYGVSEVTTK 180
QY 204 MLCADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHFLP 263
DB 181 MLCADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHFLP 240
QY 264 IRSHTKEENGLAL 276
DB 241 IRSHTKEENGLAL 253

RESULT 20
US-08-560-098A-49
; Sequence 49, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENDET, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
```


SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/07/942,157A
APPLICATION NUMBER: US/07/942,157A
FILING DATE: 19920908
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/631,673
FILING DATE: 20-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: TS1108Cont.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)815-6508
TELEFAX: (404)815-6555
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..19
OTHER INFORMATION: /label= peptide
OTHER INFORMATION: /note= "WAP signal"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 198..203
OTHER INFORMATION: /label= modified
OTHER INFORMATION: /note= "six amino acids deleted in mutant"
US-07-942-157A-3

Query Match 99.8%; Score 1505; DB 1; Length 430;
Best Local Similarity 99.6%; Pred. No. 7.8e-148;
Matches 275; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSPPEELKFCQCKQLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVYVCGSLIS 60
DB 155 KPSPPEELKFCQCKQLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVYVCGSLMS 214
QY 61 PCWISATHCFIDYPKKEDYVILGRSLNSNTQGMKFVENILHKDYSADTLAHND 120
DB 215 PCWISATHCFIDYPKKEDYVILGRSLNSNTQGMKFVENILHKDYSADTLAHND 274
QY 121 IALLKIRSGRCQAPSTIQTICLPSMYNDPQGTSCITGFKENSTDYLYPEQLKMT 180
DB 275 IALLKIRSGRCQAPSTIQTICLPSMYNDPQGTSCITGFKENSTDYLYPEQLKMT 334
QY 181 VVKLSHRECOQPHYGSEVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 335 VVKLSHRECOQPHYGSEVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVS 394
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 395 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 430

RESULT 16
US-08-560-098A-45
Sequence 45, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700

CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 306 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-45

Query Match 99.2%; Score 1496; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 4.1e-147; Indels 0; Gaps 0;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SSPPEELKFCQCKQLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVYVCGSLISPC 62
DB 33 SSPPEELKFCQCKQLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVYVCGSLISPC 92
QY 63 WVISATHCFIDYPKKEDYVILGRSLNSNTQGMKFVENILHKDYSADTLAHNDIA 122
DB 93 WVISATHCFIDYPKKEDYVILGRSLNSNTQGMKFVENILHKDYSADTLAHNDIA 152
QY 123 LKIRSGRCQAPSTIQTICLPSMYNDPQGTSCITGFKENSTDYLYPEQLKMTV 182
DB 153 LKIRSGRCQAPSTIQTICLPSMYNDPQGTSCITGFKENSTDYLYPEQLKMTV 212
QY 183 KLISHRECOQPHYGSEVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVSWG 242
DB 213 KLISHRECOQPHYGSEVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVSWG 272
QY 243 RGCAALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 273 RGCAALKDKPGVYTRVSHFLPWIRSHTKENGLAL 306

RESULT 17
US-08-560-098A-46
Sequence 46, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA

; Sequence 47, Application US/08560098A

; Patent No. 5976841

; GENERAL INFORMATION:

; APPLICANT: WENDET, Stephan

; APPLICANT: HEINZEL-WIELAND, Regina

; APPLICANT: STEFFENS, Gerd Josef

; TITLE OF INVENTION: Proteins having Fibrinolytic and

; TITLE OF INVENTION: Coagulation-inhibiting Properties

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan

; STREET: 1200 G Street, N.W., Suite 700

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30 (BPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/560,098A

; FILING DATE: 17-NOV-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: P 44 40 892.7

; FILING DATE: 17-NOV-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: EVANS, Joseph D.

; REGISTRATION NUMBER: 26,269

; REFERENCE/DOCKET NUMBER: 148/42448

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 628-8800

; TELEFAX: (202) 628-8844

; INFORMATION FOR SEQ ID NO: 47:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 432 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-560-098A-47

Query Match 100.0%; Score 1508; DB 2; Length 432;

Best Local Similarity 100.0%; Pred. No. 3.8e-148;

Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPELKFQCGQKTLRPFKLIIGGEFTTIENQPFPAIYRRHGGSVTVYCGGSLIS 60

Db 157 KPSSPPELKFQCGQKTLRPFKLIIGGEFTTIENQPFPAIYRRHGGSVTVYCGGSLIS 216

Qy 61 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGENKFEVENLILHKDYSADTLAHND 120

Db 217 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGENKFEVENLILHKDYSADTLAHND 276

Qy 121 IALLKIRSGRCQAQPSRTIQTICLPSMYNDPFGTSCETITGFGKENSTDYLYPEQLKMT 180

Db 277 IALLKIRSGRCQAQPSRTIQTICLPSMYNDPFGTSCETITGFGKENSTDYLYPEQLKMT 336

Qy 181 VVKLSHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240

Db 337 VVKLSHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 396

Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276

Db 397 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 432

RESULT 14

US-09-403-736-2

; Sequence 2, Application US/09403736

; Patent No. 6638502

; GENERAL INFORMATION:

; APPLICANT: Aventis S.A.

; APPLICANT: LI, Hong

; APPLICANT: LU, He

; APPLICANT: GRISCELLI, Frank

; APPLICANT: OPOLO, Paule

; APPLICANT: SORIA, Claudine

; APPLICANT: RAGOT, Thierry

; APPLICANT: LEGRAND, Yves

; APPLICANT: SORIA, Jeanette

; APPLICANT: MABILAT, Christelle

; APPLICANT: PERRICAUDET, Michel

; APPLICANT: YEH, Patrice

; TITLE OF INVENTION: Adenovirus-Mediated Intratumoral Delivery Of An Angioge

; TITLE OF INVENTION: For The Treatment of Tumors

; FILE REFERENCE: A2778A-US

; CURRENT APPLICATION NUMBER: US/09/403,736

; CURRENT FILING DATE: 1999-10-26

; PRIOR APPLICATION NUMBER: PCT/EP98/02491

; PRIOR FILING DATE: 1998-11-05

; PRIOR APPLICATION NUMBER: 60/044,980

; PRIOR FILING DATE: 1997-04-28

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 411

; TYPE: PRT

; ORGANISM: humanurokinase

; US-09-403-736-2

Query Match 99.8%; Score 1505; DB 4; Length 411;

Best Local Similarity 99.6%; Pred. No. 7.3e-148;

Matches 275; Conservative 1; Mismatches 0; Indels 0; Gaps 0

Qy 1 KPSSPPELKFQCGQKTLRPFKLIIGGEFTTIENQPFPAIYRRHGGSVTVYCGGSLIS 60

Db 136 KPSSPPELKFQCGQKTLRPFKLIIGGEFTTIENQPFPAIYRRHGGSVTVYCGGSLIS 19:

Qy 61 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGENKFEVENLILHKDYSADTLAHND 120

Db 196 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGENKFEVENLILHKDYSADTLAHND 25:

Qy 121 IALLKIRSGRCQAQPSRTIQTICLPSMYNDPFGTSCETITGFGKENSTDYLYPEQLKMT 180

Db 256 IALLKIRSGRCQAQPSRTIQTICLPSMYNDPFGTSCETITGFGKENSTDYLYPEQLKMT 31:

Qy 181 VVKLSHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 24:

Db 316 VVKLSHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 37:

Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276

Db 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 15

US-07-942-157A-3

; Sequence 3, Application US/07942157A

; Patent No. 5648253

; GENERAL INFORMATION:

; APPLICANT: Wei, Cha-Mer

; TITLE OF INVENTION: Inhibitor-Resistant Urokinase

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Kilpatrick & Cody

; STREET: 1100 Peachtree Street Suite 2800

; CITY: Atlanta

; STATE: Georgia

; COUNTRY: USA

; ZIP: 30309-4530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

Db 256 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFSGTSCITGFGKENSIDYLYPEQLKMT 315
Qy 181 VVKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db 316 VVKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
Db 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 10
5219569-2
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO:2
; LENGTH: 430
5219569-2

Query Match 100.0%; Score 1508; DB 6; Length 430;
Best Local Similarity 100.0%; Pred. No. 3.8e-148; Indels 0; Gaps 0;
Matches 276; Conservative 0; Mismatches 0;
Qy 1 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGGSVTVYVCGSLIS 60
Db 155 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGGSVTVYVCGSLIS 214
Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
Db 215 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 274
Qy 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFSGTSCITGFGKENSIDYLYPEQLKMT 180
Db 275 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFSGTSCITGFGKENSIDYLYPEQLKMT 334
Qy 181 VVKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db 335 VVKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 394
Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
Db 395 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 430

RESULT 11
US-09-101-272G-1
; Sequence 1, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: mat_peptide
; LOCATION: (21)..
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (20)..
; OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)
US-09-101-272G-1

Query Match 100.0%; Score 1508; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.8e-148; Indels 0; Gaps 0;
Matches 276; Conservative 0; Mismatches 0;
Qy 1 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGGSVTVYVCGSLIS 60
Db 156 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGGSVTVYVCGSLIS 215
Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
Db 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275
Qy 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFSGTSCITGFGKENSIDYLYPEQLKMT 180
Db 276 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFSGTSCITGFGKENSIDYLYPEQLKMT 335
Qy 181 VVKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db 336 VVKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
Db 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 12
5188829-1
; Patent No. 5188829
; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEKI; YAMADA, CHIKAKO
; TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1988
; SEQ ID NO:1
; LENGTH: 431
5188829-1

Query Match 100.0%; Score 1508; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.8e-148; Indels 0; Gaps 0;
Matches 276; Conservative 0; Mismatches 0;
Qy 1 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGGSVTVYVCGSLIS 60
Db 156 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGGSVTVYVCGSLIS 215
Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
Db 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275
Qy 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFSGTSCITGFGKENSIDYLYPEQLKMT 180
Db 276 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFSGTSCITGFGKENSIDYLYPEQLKMT 335
Qy 181 VVKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db 336 VVKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
Db 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 13
US-08-560-098A-47

NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 92H832
TELEPHONE: (908) 665 2400
TELEFAX: (908) 771 6159
TELEX: 219484
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-153-799-18

Query Match 100.0%; Score 1508; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 3.6e-148;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGGSLIS 60
DB 136 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGGSLIS 195
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLILHKDYSADTLAHND 120
DB 196 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLILHKDYSADTLAHND 255
QY 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 180
DB 256 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 315
QY 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 316 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 9

US-09-181-816-1
Sequence 1, Application US/09181816
Patent No. 6277818
GENERAL INFORMATION:
APPLICANT: MAZAR, Andrew P.
APPLICANT: JONES, Terence R.
TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR
FILE REFERENCE: 329042000300 SIDN 1-7
CURRENT APPLICATION NUMBER: US/09/181,816
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 411
TYPE: PRT
ORGANISM: Homo sapiens
US-09-181-816-1

Query Match 100.0%; Score 1508; DB 3; Length 411;
Best Local Similarity 100.0%; Pred. No. 3.6e-148;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGGSLIS 60
DB 136 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGGSLIS 195
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLILHKDYSADTLAHND 120
DB 196 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLILHKDYSADTLAHND 255
QY 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 180

TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-286-748B-18

Query Match 100.0%; Score 1508; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 3.6e-148;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGGSLIS 60
DB 136 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGGSLIS 195
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLILHKDYSADTLAHND 120
DB 196 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLILHKDYSADTLAHND 255
QY 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 180
DB 256 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 315
QY 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 316 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 8

US-08-153-799-18
Sequence 18, Application US/08153799
Patent No. 5766883
GENERAL INFORMATION:
APPLICANT: Ballance, David J
APPLICANT: Goodey, Andrew R
TITLE OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESS: R Hain Swope, BOC Health Care Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill
STATE: New Jersey
COUNTRY: USA
ZIP: 07974
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650
FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:

```
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 100.0%; Score 1508; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 3.3e-148;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFALYYRRHGGSVTVVCGGSLIS 60
DB 91 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFALYYRRHGGSVTVVCGGSLIS 150
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKOYSADTLAHND 120
DB 151 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKOYSADTLAHND 210
QY 121 IALLKIRSKGRCQAPSRITQICLPWMYNDPQGTSCETITGFGKENSTDYLYPEQLKWT 180
DB 211 IALLKIRSKGRCQAPSRITQICLPWMYNDPQGTSCETITGFGKENSTDYLYPEQLKWT 270
QY 181 VVKLISHRECQPHYGVSEVTTKMLCAADPWKTDSCQSGGGLVCSLQGRMTLTGIVS 240
DB 271 VVKLISHRECQPHYGVSEVTTKMLCAADPWKTDSCQSGGGLVCSLQGRMTLTGIVS 330
QY 241 WGRGALKDKPGVTVRVSHFLPWIRSHTKENGLAL 276
DB 331 WGRGALKDKPGVTVRVSHFLPWIRSHTKENGLAL 366

RESULT 6
US-08-087-163-1
; Sequence 1, Application US/08087163
; Patent No. 5472692
; GENERAL INFORMATION:
; APPLICANT: Liu, Jian-Ning
; TITLE OF INVENTION: PRO-UKINASE MUTANTS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,163
; FILING DATE: 07/02/93
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983

; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
US-08-087-163-1

Query Match 100.0%; Score 1508; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 3.6e-148;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFALYYRRHGGSVTVVCGGSLIS 60
DB 136 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFALYYRRHGGSVTVVCGGSLIS 195
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKOYSADTLAHND 120
DB 196 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKOYSADTLAHND 255
QY 121 IALLKIRSKGRCQAPSRITQICLPWMYNDPQGTSCETITGFGKENSTDYLYPEQLKWT 180
DB 256 IALLKIRSKGRCQAPSRITQICLPWMYNDPQGTSCETITGFGKENSTDYLYPEQLKWT 315
QY 181 VVKLISHRECQPHYGVSEVTTKMLCAADPWKTDSCQSGGGLVCSLQGRMTLTGIVS 240
DB 316 VVKLISHRECQPHYGVSEVTTKMLCAADPWKTDSCQSGGGLVCSLQGRMTLTGIVS 375
QY 241 WGRGALKDKPGVTVRVSHFLPWIRSHTKENGLAL 276
DB 376 WGRGALKDKPGVTVRVSHFLPWIRSHTKENGLAL 411

RESULT 7
US-08-286-748B-18
; Sequence 18, Application US/08286748B
; Patent No. 5759542
; GENERAL INFORMATION:
; APPLICANT: Victor Gurewich
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
; TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,748B
; FILING DATE: August 5, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: J. Peter Fasse
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04547/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
```

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-44

Query Match 100.0%; Score 1508; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 3.3e-148;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVYCGSLIS 60
DB 91 KPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVYCGSLIS 150
QY 61 PCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHND 120
DB 151 PCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHND 210
QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGKENSTDYLYPEQLKMT 180
DB 211 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGKENSTDYLYPEQLKMT 270
QY 181 VKLISHRECOQPHYGVSEVTTKMLCAADPWKMTDCSQGSDGGLVCSLQGRMTLTGIVS 240
DB 271 VKLISHRECOQPHYGVSEVTTKMLCAADPWKMTDCSQGSDGGLVCSLQGRMTLTGIVS 330
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGIAL 276
DB 331 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGIAL 366

RESULT 4
US-08-967-024C-24
Sequence 24, Application US/08967024C
Patent No. 6133011
GENERAL INFORMATION:
APPLICANT: WNEEDT, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOSHA, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967.024C
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-24
Query Match 100.0%; Score 1508; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 3.3e-148;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVYCGSLIS 60
DB 91 KPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVYCGSLIS 150
QY 61 PCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHND 120
DB 151 PCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHND 210
QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGKENSTDYLYPEQLKMT 180
DB 211 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGKENSTDYLYPEQLKMT 270
QY 181 VKLISHRECOQPHYGVSEVTTKMLCAADPWKMTDCSQGSDGGLVCSLQGRMTLTGIVS 240
DB 271 VKLISHRECOQPHYGVSEVTTKMLCAADPWKMTDCSQGSDGGLVCSLQGRMTLTGIVS 330
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGIAL 276
DB 331 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGIAL 366

RESULT 5
US-08-967-024C-25
Sequence 25, Application US/08967024C
Patent No. 6133011
GENERAL INFORMATION:
APPLICANT: WNEEDT, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOSHA, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967.024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8

	PROPEP	22	33	BY SIMILARITY.
F	CHAIN	34	566	TISSUE-TYPE PLASMINOGEN ACTIVATOR.
F	CHAIN	34	314	TISSUE-TYPE PLASMINOGEN ACTIVATOR A CHAIN.
F	CHAIN	315	566	TISSUE-TYPE PLASMINOGEN ACTIVATOR B CHAIN.
F	DOMAIN	40	82	PIRONECTIN TYPE-I.
F	DOMAIN	83	121	EGF-LIKE.
F	DOMAIN	128	209	KRINGLE 1.
F	DOMAIN	219	300	KRINGLE 2.
F	DOMAIN	315	566	SERINE PROTEASE.
F	ACT SITE	361	361	CHARGE RELAY SYSTEM.
F	ACT_SITE	410	410	CHARGE RELAY SYSTEM.
F	ACT_SITE	517	517	CHARGE RELAY SYSTEM.
F	DISULFID	42	72	BY SIMILARITY.
F	DISULFID	70	79	BY SIMILARITY.
F	DISULFID	87	98	BY SIMILARITY.
F	DISULFID	92	109	BY SIMILARITY.
F	DISULFID	111	120	BY SIMILARITY.
F	DISULFID	128	209	BY SIMILARITY.
F	DISULFID	149	191	BY SIMILARITY.
F	DISULFID	180	204	BY SIMILARITY.
F	DISULFID	219	300	BY SIMILARITY.
F	DISULFID	240	282	BY SIMILARITY.
F	DISULFID	271	295	BY SIMILARITY.
F	DISULFID	303	434	INTERCHAIN (BY SIMILARITY).
F	DISULFID	346	362	BY SIMILARITY.
F	DISULFID	354	423	BY SIMILARITY.
F	DISULFID	448	523	BY SIMILARITY.
F	DISULFID	480	496	BY SIMILARITY.
F	DISULFID	513	541	BY SIMILARITY.
F	CARBOHYD	153	153	N-LINKED (GLCNAC. .) (POTENTIAL).
F	CARBOHYD	487	487	N-LINKED (GLCNAC. .) (POTENTIAL).
F	SEQUENCE	566 AA;	63701 MW;	2EB56BE4E32276C3 CRC64;
uery Match		36.1%;	Score 814;	DB 1; Length 566;
eat Local Similarity		35.8%;	Pred. No. 4.7e-58;	
atches 176;	Conservative	60;	Mismatches 155;	Indels 100; Gaps 9;
C	3	ELH0VP-SNCD---	CLNGSGTCVSNKYF5NIHWCNPKFGGQHCIDKSKTCYEGNGHFY	58
I	78	QCHSVPRSRSEPCWFCFNGGTCRQALYSSPF-VQCPEGPMKLCIEDATATCYKDOGVAY	136	
C	59	RGKASTDTMGRPCLPWNSATVLOQTYAHRSDALQGLGKHNYCRNPDRRPRPCYVQVG	118	
I	137	RGTWSTAESGAECANWNSGLAMKPYSGRRPNAIRLGLGNHNYCRNPQDQSKPCYVYFKA	196	
C	119	LKPLVQBCMHDCR-----	132	
I	197	GKYISEFCSTPACAKVAEEDGDCYTNGNLAYRGTSHTKSGASCLPWN5VFLTSKIITAW	256	
C	133	-----DGKLRFPQ-----	CG-QKTLRPRFKII 152	
I	257	KSNAPALGLGKHNCRPDGDAPQWCHVWKQRQLTWEYCDVPQCVTCGLRQYKEPQPRIK	316	
C	153	GGEFTTENQWPAALY-RRHRGGSVTYVCGSGLISPCWVTSATHCFIDYPKKEDIYVYL	211	
I	317	GGLEFADITSHPWQAAIFVQVRRSPGRFLCGGILISSCWVL5AAHCFQERYPPPHLKVFL	376	
C	212	GRSLRNSNTQCEMKFEVENILHKDYSADT'LAHNDIALLKIRSGEGRCAOPSTIOTIC	271	
I	377	GRVTLVPGEEEQTFVEVKI'IIHKFDDDT--YDNDIALHLK5DLSLTCASESASVETIC	434	
C	272	LPSMYNDPQFQTSCEITFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKM	331	
I	435	LPDASLQPDWTECELSGYKGKHSSPPFFSERLKEARVELYPSRSTCSQHLFNRTVTNNM	494	
C	332	LCAADPQW-----KTPSCGDSGGPLVCSLQGRWMLTGI5VSGRGCAKDKPGVYTVFS	385	
I	495	LCAGDTRGGGHTNLHDACQDGGGPLVCMKNHMTLVGLISWGJCGGRKDPGVYTKVT	554	
C	386	HFLPWIRSHTX 396	---	---

```

DB      555 NYLDWIRNTR 565

RESULT 15
URTG_DESO
AC      URTG_DESRO      STANDARD;      PRT;      394 AA.
AC      P49150;
DT      01-FEB-1996 (Rel. 33, Created)
DT      01-FEB-1996 (Rel. 33, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA
DE      gamma).
DE      Desmodus rotundus (Vampire bat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC      Desmodontinae; Desmodus.
OX      NCBI_TaxID=9430;
RN      [1]
RN      SEQUENCE FROM N.A.
RN      TISSUE=Salivary gland;
RX      MEDLINE=92039036; PubMed=1337019;
RX      Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA      Alagon A., Donner P., Schleuning W.D.;
RT      "The plasminogen activator family from the salivary gland of the
RT      vampire bat Desmodus rotundus: cloning and expression.";
RL      Gene 105:229-237(1991).
RN      [2]
RN      CHARACTERIZATION.
RN      MEDLINE=93353059; PubMed=1309059;
RX      Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA      Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA      Donner P.;
RT      "Plasminogen activators from the saliva of Desmodus rotundus (common
RT      vampire bat): unique fibrin specificity.";
RN      Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC      -!- FUNCTION: Probably essential to support the feeding habits of this
CC      exclusively haematophagous animal. Probable potent thrombolytic
CC      agent.
CC      -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC      plasminogen to form plasmin.
CC      -!- SUBUNIT: Monomer.
CC      -!- SIMILARITY: Belongs to peptidase family S1.
CC      -!- SIMILARITY: Contains 1 kringle domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M63990; AAA31595.1; -.
DR      PIR; J50600; JS0600.
DR      HSSP; P98119; 1A51.
DR      MEROPS; S01.239; -.
DR      InterPro; IPR009003; Cys_Ser_trypsin.
DR      InterPro; IPR000001; Kringle.
DR      InterPro; IPR001254; Peptidase S1.
DR      InterPro; IPR001314; Peptidase_S1A.
DR      Pfam; PF000051; kringle; 1.
DR      Pfam; PF00089; trypsin; 1.
DR      PRINTS; PR00722; CHYMOTRYPSIN.
DR      PRINTS; PR00015; KRINGLE.
DR      ProDom; PD000395; Kringle; 1.
DR      SMART; SM00130; KR; 1.
DR      SMART; SM00020; Tryp_SPC; 1.
DR      PROSITE; PS00021; KRINGLE_1; 1.
DR      PROSITE; PS00070; KRINGLE_2; 1.
DR      PROSITE; PS00240; TRYPsin DOM; 1.
DR      PROSITE; PS00134; TRYPsin HIS; 1.
DR      PROSITE; PS00135; TRYPsin SER; 1.
DR      Plasminogen activator; Hydrolase; Serine protease; Glycoprotein;
KW

```


Kringle; Signal; Multigene family.
 RT SIGNAL 1 36 POTENTIAL
 RT CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
 RT DOMAIN 45 126 KRINGLE.
 RT ACT_SITE 142 394 SERINE PROTEASE.
 RT ACT_SITE 189 394 CHARGE RELAY SYSTEM (BY SIMILARITY).
 RT ACT_SITE 238 394 CHARGE RELAY SYSTEM (BY SIMILARITY).
 RT ACT_SITE 345 394 CHARGE RELAY SYSTEM (BY SIMILARITY).
 RT DISULFID 45 126 BY SIMILARITY.
 RT DISULFID 66 108 BY SIMILARITY.
 RT DISULFID 97 121 BY SIMILARITY.
 RT DISULFID 131 262 BY SIMILARITY.
 RT DISULFID 174 190 BY SIMILARITY.
 RT DISULFID 262 251 BY SIMILARITY.
 RT DISULFID 276 351 BY SIMILARITY.
 RT DISULFID 308 324 BY SIMILARITY.
 RT DISULFID 341 369 BY SIMILARITY.
 RT CARBOHYD 315 315 N-LINKED (GLNAC...) (POTENTIAL).
 SEQUENCE 394 AA; 44105 MW; 9CCDEF52F3D81FCD CRC64;
 ary Match 33.5%; Score 756; DB 1; Length 394;
 at Local Similarity 42.9%; Pred. No. 1.4e-53;
 tches 155; Conservative 56; Mismatches 134; Indels 16; Gaps 7;
 Qy 45 DKSCTCYEGNGHFFRGKASTDTWGRPCLPNSATVLOQTYHAHRSALQGLGKHNVCN 104
 Dc 40 DPHATCYKDGQVTVRGTWSTSESQAQCINNNSLLIRTYNGRMPKAVLGLGNHNYCRN 99
 Qy 105 PDNRERRWCYVQGLKPLVQSCMVDHCDAGKLFQCG-QKTLRPRFKIIGGEFTTIENQP 163
 Dc 100 PDGASKPWCYVVKARKFTSSECSVPVCS---KATCGLRKYKPEQLHSTGGLFTDITSH 155
 Qy 164 WFAIYRHRGGS-VTVYCGSLISPCWVISAHCYFD-YPKEDYIVYLGSRSLNSNTQ 221
 Dc 156 WQAAIFAQNRSSGGERFLCGILISSCWLTAAHCFQERYPPQHLRVV-LGRTRYVKPGK 214
 Qy 222 GEMKFEVENILTKHYSADTLAHNDIALKIRSKGRCQAPSRITQICLPMSYNDPQF 281
 Dc 215 EQQTFVEKCIHVEFDQDT--YNDIALQLKSGSPQCAQSDSVRAICLPEANILQPD 272
 Qy 282 GTSCEITGFGKENTDLYPEQLKMTVVKULISHRECQPHYGVSEVITKMLCAADQWKT 341
 Dc 273 WTECELSGYGKHSSPFYSQLEKGVHRLVYPSRSCTKFLFNKVTNNMLCAGDTRSGE 332
 Qy 342 -----DSCQDSSGGLVCSLQGRMTLTGIVSWGRGCALDKPGVYVRVSHFLPWTRSH 395
 Dc 333 IYPNVHDACQDSSGGLVCSYNDNEMTLGLISWGVGCGEKDIFGVYTKVTNYLGWIRDNM 392
 Qy 396 K 396
 Dc 393 R 393
 RE LT 16
 HG HUMAN
 ID HGFA_HUMAN STANDARD; PRT; 655 AA.
 AC Q04756; Q14726;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DI 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
 Dc activator) (HGFA).
 GN HGFA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RI SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Liver, and Serum;
 RX MEDLINE=93252878; PubMed=7663665;
 RP Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
 RA Kitamura N.;

"Molecular cloning and sequence analysis of the cDNA for a human
 serine protease responsible for activation of hepatocyte growth
 factor. Structural similarity of the protease precursor to blood
 coagulation factor XII.".
 J. Biol. Chem. 268:10024-10028(1993).
 [2]
 SEQUENCE OF 40-655 FROM N.A.
 Zhao S., Odell C.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 CC FUNCTION: Activates hepatocyte growth factor (HGF) by
 converting it from a single chain to a heterodimeric form.
 CC SUBUNIT: Dimer of a short chain and a long chain linked by a
 disulfide bond.
 CC SUBCELLULAR LOCATION: Secreted as an inactive single-chain
 precursor and is then activated to a heterodimeric form.
 CC TISSUE SPECIFICITY: Liver.
 CC SIMILARITY: Belongs to peptidase family S1.
 CC SIMILARITY: Contains 2 EGF-like domains.
 CC SIMILARITY: Contains 1 fibronectin type I domain.
 CC SIMILARITY: Contains 1 fibronectin type II domain.
 CC SIMILARITY: Contains 1 kringle domain.
 CC CAUTION: It is uncertain whether Met-1 is the initiator.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 or send an email to license@isb-sib.ch).

 EMBL; D14012; BRA03113.1; --
 EMBL; Z69923; -- NOT_ANNOTATED_CDS.
 PIR; A46688; A46688.
 HSRP; P00763; IDPO.
 MEROFS; S01.228; --
 Genew; HGNC:4894; HGFAC.
 MIM; 604552; --
 GO; GO:0005576; C:extracellular; TAS.
 GO; GO:0004253; P:serine-type endopeptidase activity; TAS.
 GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
 InterPro; IPR009003; Cys_Ser_trypsin.
 InterPro; IPR000742; EGF_2.
 InterPro; IPR006209; EGF-like.
 InterPro; IPR000083; Fibrinctn1.
 InterPro; IPR000562; FN_Type_II.
 InterPro; IPR006210; IEGF.
 InterPro; IPR000001; Kringle.
 InterPro; IPR001254; Peptidase_S1.
 InterPro; IPR001314; Peptidase_S1A.
 Pfam; PF00008; EGF_2.
 Pfam; PF00039; fn1; 1.
 Pfam; PF00040; fn2; 1.
 Pfam; PF00051; kringle; 1.
 Pfam; PF00089; trypsin; 1.
 PRINTS; PRO0722; CHYMOTRYPSIN.
 PRINTS; PRO0013; FNTYPEII.
 PRINTS; PRO0018; KRINGLE.
 ProDom; PD000995; FN_Type_II; 1.
 ProDom; PD000395; Kringle; 1.
 SMART; SM00181; EGF_2.
 SMART; SM00058; FN1; 1.
 SMART; SM00059; FN2; 1.
 SMART; SM00130; KR; 1.
 SMART; SM00020; TRYP_SPC; 1.
 PROSITE; PS00022; EGF_1; 2.
 PROSITE; PS01186; EGF_2; 1.
 PROSITE; PS00026; EGF_3; 2.
 PROSITE; PS01253; FIBRONECTIN_1; 1.
 PROSITE; PS00023; FIBRONECTIN_2; 1.
 PROSITE; PS00021; KRINGLE_1; 1.
 PROSITE; PS00070; KRINGLE_2; 1.
 PROSITE; PS00240; TRYPSIN_DOM; 1.

I PROSITE; PS00134; TRYPSIN HIS; 1.
I PROSITE; PS00135; TRYPSIN SER; 1.
I Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
I EGF-like domain; Repeat; Zymogen.
I SIGNAL 1 30
I PROPEP 31 372
I CHAIN 373 407
I
I CLEAVED IN ACTIVE FORM.
I HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
I CHAIN.
I CHAIN 408 655
I
I DOMAIN 108 148
I EGF-LIKE 1.
I FIBRONECTIN TYPE-I.
I DOMAIN 200 240
I EGF-LIKE 2.
I DOMAIN 241 279
I KRINGLE.
I DOMAIN 286 365
I SERINE PROTEASE.
I ACT_SITE 447 447
I CHARGE RELAY SYSTEM (BY SIMILARITY).
I ACT_SITE 497 497
I CHARGE RELAY SYSTEM (BY SIMILARITY).
I ACT_SITE 598 598
I DISULFID 108 133
I BY SIMILARITY.
I DISULFID 122 148
I BY SIMILARITY.
I DISULFID 164 175
I BY SIMILARITY.
I DISULFID 169 186
I BY SIMILARITY.
I DISULFID 188 197
I BY SIMILARITY.
I DISULFID 202 230
I BY SIMILARITY.
I DISULFID 228 237
I BY SIMILARITY.
I DISULFID 245 256
I BY SIMILARITY.
I DISULFID 250 267
I BY SIMILARITY.
I DISULFID 269 278
I BY SIMILARITY.
I DISULFID 286 367
I BY SIMILARITY.
I DISULFID 307 349
I BY SIMILARITY.
I DISULFID 338 362
I INTERCHAIN (BY SIMILARITY).
I DISULFID 394 521
I BY SIMILARITY.
I DISULFID 432 448
I BY SIMILARITY.
I DISULFID 440 510
I BY SIMILARITY.
I DISULFID 535 604
I BY SIMILARITY.
I DISULFID 567 583
I BY SIMILARITY.
I DISULFID 594 622
I BY SIMILARITY.
I N-LINKED (GLCNAC. . .) (POTENTIAL).
I CARBOHYD 48 48
I N-LINKED (GLCNAC. . .) (POTENTIAL).
I CARBOHYD 290 290
I N-LINKED (GLCNAC. . .) (POTENTIAL).
I CARBOHYD 468 468
I N-LINKED (GLCNAC. . .) (POTENTIAL).
I CARBOHYD 492 492
I N-LINKED (GLCNAC. . .) (POTENTIAL).
I CARBOHYD 546 546
I N-LINKED (GLCNAC. . .) (POTENTIAL).
I CONFLICT 644 644
I R -> Q (IN REF. 2).
I SEQUENCE 655 AA; 70681 MW; 2CF72F1B1B862ED7 CRC64;
I
I very Match 32.2%; Score 726.5; DB 1; Length 655;
I est Local Similarity 37.2%; Pred. No. fe-51;
I atches 158; Conservative 53; Mismatches 161; Indels 53; Gaps 9;
I
I 5 HQVPSNCDLNGGTCVSNKYFSNTHW-----CNCPKKFGGOHCEDKSKTCYEGNGH 56
I 242 HTACLSFPLNGGTC-----HLIVATGTTVCACPPGFAGKLCNIEPDERCFLNGT 292
I
I 57 FYRGKASTDTMGRPLPNSATVLCQTYHAHRSDALQGLGKHNCRPNRRRPWCYV- 115
I 293 GYRGVASTSAGLSCLAWSNLLYQELHVDVSGAALLGLGPHAYCRNPDRPWCYV 352
I
I 116 -----QVGLKPLVQECMVHDCADGKLKFCQGX-----TLRPFKI 151
I 353 KDSALSWEYCYKLEACESTRVLQSPDLATLPEPASPGQ--ACGRHKKRTFLRPR--I 408
I
I 152 IGGFTTTIENQWFAAIYRRHRGGSVTYVCGSLISPCWISATHCFIDYPKKEDITYVL 211
I 409 IGGSSSLPGSHPLAAIY---IGDS---FCAGSLVHTCWVSAACFSPHSPRDSVSWL 462
I
I 212 GRSRLNSNTQEMKFEVENLILHKDYSADTLAHNDIALKIRSGECRQAPSTRTIOTIC 271
I 463 GQHPFNRTTDTVTQFGIEKIPYTLYSVFNPSDSH-DLVILRLKKKGDRCATRSQVQPIC 521
I
I 272 LPSMYNDPQFQTSBITGFGKENGSTDYLPQLKMTVKLISHRECOQPHYGVSEVTTKM 331
I 522 LPEGSTFPAGHKQIAGWGHLDENVSGYSSSLREALVPLVADHKCSSPEVYAGDISPNM 591

QY 332 LCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSMRGKCALDKPKGVYTRVSHFLPMI 391
Db 582 LCRGFFDCKSDACQDGGPLACRKNVAYLYGLISWGDGCGRLHKPKGVYTRVANYVDWI 641
QY 392 RSHTX 396
Db 642 NDRIR 646
RESULT 17
HGFA MOUSE
ID HGFA MOUSE STANDARD; PRT; 653 AA.
AC Q9RQ98; Q9JKV4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/c;
RA Itoh H., Kataoka H., Koono H.;
RT "Mouse hepatocyte growth factor activator."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21226753; PubMed=11032833;
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J.,
RA Yang J., Huan Y.;
RT "Activation of hepatocyte growth factor (HGF) by endogenous HGF
RT activator is required for metanephric kidney morphogenesis in
RT vitro";
RL J. Biol. Chem. 276:15099-15106(2001).
CC -!- FUNCTION: Activates hepatocyte growth factor (HGF) by converting
CC it from a single chain to a heterodimeric form (By similarity).
CC -!- SUBUNIT: Dimer of a short chain and a long chain linked by a
CC disulfide bond (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted as an inactive single-chain
CC precursor and is then activated to a heterodimeric form (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF099017; AAF02489.1; -
DR EMBL; AF224724; AAF34712.1; -
DR HSSP; P00763; LDPO.
DR MEROPS; S01.238; -
DR MGD; MGI:1859281; Hgfac.
DR InterPro; IPR009003; Cys_ser_trypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.

InterPro: IPR001314; Peptidase_SIA.
 Pfam: PF00008; EGF; 2.
 Pfam: PF00039; En1; 1.
 Pfam: PF00040; En2; 1.
 Pfam: PF00051; kringle; 1.
 Pfam: PF00089; trypsin; 1.
 PRINTS: PR00722; CHYMOTRYPSIN.
 PRINTS: PR00013; ENTPPEI.
 PRINTS: PR00018; KRINGLE.
 ProDom: PD000395; FN_type_II; 1.
 ProDom: PD000395; Kringle; 1.
 SMART: SM00181; EGF; 2.
 SMART: SM00059; FN2; 1.
 SMART: SM00130; KR; 1.
 SMART: SM00020; Tryp_spc; 1.
 PROSITE: PS00022; EGF_1; 2.
 PROSITE: PS01186; EGF_2; 1.
 PROSITE: PS00026; EGF_3; 2.
 PROSITE: PS01253; FIBRONECTIN_1; 1.
 PROSITE: PS00023; FIBRONECTIN_2; 1.
 PROSITE: PS00021; KRINGLE_1; 1.
 PROSITE: PS00070; KRINGLE_2; 1.
 PROSITE: PS50240; TRYPSIN_DOM; 1.
 PROSITE: PS00134; TRYPSIN_HIS; 1.
 PROSITE: PS00135; TRYPSIN_SER; 1.
 K0 Hydroxylase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
 K0 EGF-like domain; Repeat; Zymogen.
 F1 SIGNAL 1 29
 F1 PROPEP 30 369
 F1 CHAIN 370 405
 F1 CHAIN 406 653
 F1 DOMAIN 105 145
 F1 DOMAIN 157 195
 F1 DOMAIN 197 237
 F1 DOMAIN 238 276
 F1 DOMAIN 283 364
 F1 DOMAIN 406 653
 F1 ACT_SITE 445 445
 F1 ACT_SITE 495 495
 F1 ACT_SITE 596 596
 F1 DISULFID 105 130
 F1 DISULFID 119 145
 F1 DISULFID 161 172
 F1 DISULFID 166 183
 F1 DISULFID 185 194
 F1 DISULFID 199 227
 F1 DISULFID 225 234
 F1 DISULFID 242 253
 F1 DISULFID 247 264
 F1 DISULFID 266 275
 F1 DISULFID 283 364
 F1 DISULFID 304 346
 F1 DISULFID 335 359
 F1 DISULFID 392 519
 F1 DISULFID 430 446
 F1 DISULFID 438 508
 F1 DISULFID 533 602
 F1 DISULFID 565 581
 F1 DISULFID 592 620
 F1 CARBOHYD 39 39
 F1 CARBOHYD 47 47
 F1 CARBOHYD 63 63
 F1 CARBOHYD 287 287
 F1 CARBOHYD 466 466
 F1 CARBOHYD 544 544
 F1 CONFLICT 164 164
 F1 SEQUENCE 653 AB; 70567 NW; 8864B2025DF7FDC CRC64;
 ery Match 31.7%; Score 716; DB 1; Length 653;
 st Local Similarity 36.3%; Pred. NO. 4.2e-50;
 tches 154; Conservative 58; Mismatches 162; Indels 50; Gaps 8;

QY 5 HQVPSNCDCLNGSTCVSNKYFSNIHM-----CNCPKKFGGQHQEIDKSKTCYEGNGH 56
 Db 239 HTACLSPLCLNGGTC-----HLVGTGTSVCTPLGVAGRFNIVPTHEHCLNGT 289
 QY 57 FYRGKASTDTMGRECLPWN SATVLQOTYHARRSDALQGLGKHNYCRNPNRRPWCYVQ 116
 Db 290 EYRGVASTAASGLSCLAWNSDLLYOELHVDVAVALLLGLPHAYCRNPKDERPWCYV 349
 QY 117 VGLRPLVQECMVHDC-----ADGKLKFCQCGK-----TLRPRFKII 152
 Db 350 KDNALSWEYICRLTACESLARVHSOTPEILALPSPAFAVRPTCGKRHKRTFLRPR--II 407
 QY 153 GGEFTTIENQPFAAIYRRHRGSGVTVYCGSLISPCWVISATHCFIDYPKKEDIYVVLG 212
 Db 408 GGSLSLPGSHPEWLAAY---IGNS---FCAGSLVHTCWVWSAAHCFANSPRDSITVVLG 461
 QY 213 RSLNSNTQGEKMEVENLIHKDYSADTLAHHNDIALLKIRSKRSGCAQPSRTIQICL 272
 Db 462 QHFFNRITDVTQTGIEKYVPTLYSVFNPNNH-DLVLRLLKKKGCRCVRSQVQFICL 520
 QY 273 PSMYNDPQGTSCETIGFGKENSTDLYPEOLKMTVVKLISHRECQCPHYYGSEVTTML 332
 Db 521 PEAGSSPFTGHKQIAGWGHMDENVSSVNSLLEALVPLVADHKCSSPEVYGAISPML 580
 QY 333 CHADQWITDCQDGGSGPLVCSIQGRMTLTGIVSWRGCAKDKPGYITVSHFLPWIR 392
 Db 581 CAGYFDCKSDACQDGGSGPLVCEKNGVAYLYGIISWGDGCGRLNKPVGYYTVRVANYVDW 640
 QY 393 SHTK 396
 Db 641 DEIR 644
 RESULT 18
 ID FA12 CAVPO STANDARD; PRT; 603 AA.
 AC QC4962;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
 DE (HAF) (Fragment)
 GN F12.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
 RC TISSUE=Liver;
 RX MEDLINE=93003367; PubMed=1390917;
 RA Semba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,
 RA Kambara T., Okabe H.;
 RT "Primary structure of guinea-pig Hageman factor: sequence around the
 RT cleavage site differs from the human molecule.";
 RL Biochim. Biophys. Acta 1159:113-121(1992).
 CC -!- FUNCTION: Factor XII is a serum glycoprotein that participates in
 CC the initiation of blood coagulation, fibrinolysis, and the
 CC generation of bradykinin and angiotensin.
 CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
 CC VII to form factor VIIa and factor XI to form factor Xla.
 CC -!- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a
 CC complex bound to an anionic surface. Prekallikrein is cleaved by
 CC factor XII to form kallikrein, which then cleaves factor XII first
 CC to alpha-factor Xla and then to beta-factor Xlaa. Alpha-factor
 CC Xlaa activates factor XI to factor Xla.
 CC -!- SIMILARITY: Belongs to peptidase family S1.
 CC -!- SIMILARITY: Contains 2 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -!- SIMILARITY: Contains 1 kringle domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X68615; CAA48600.1; --
 PIR; S28941; S28941.
 RSP; P00763; IDPO.
 MEROPS; S01.211; --
 InterPro; IPR009003; Cys_Ser_trypsin.
 InterPro; IPR000742; EGF_2.
 InterPro; IPR001881; EGF_Ca.
 InterPro; IPR006209; EGF_like.
 InterPro; IPR000083; Fibrinctnl.
 InterPro; IPR000562; FN_type_II.
 InterPro; IPR006210; IEGF.
 InterPro; IPR000001; Kringle.
 InterPro; IPR001254; Peptidase_S1.
 InterPro; IPR001314; Peptidase_S1A.
 Pfam; PF00008; EGF; 2.
 Pfam; PF00039; fn1; 1.
 Pfam; PF00040; fn2; 1.
 Pfam; PF00051; kringle; 1.
 Pfam; PF00089; trypsin; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 PRINTS; PR00013; ENTYPBII.
 PRINTS; PR00018; KRINGLE.
 ProDom; PD000995; FN_Type_II; 1.
 ProDom; PD000395; Kringle; 1.
 SMART; SMC0181; EGF; 2.
 SMART; SMC0058; FN1; 1.
 SMART; SMC0059; FN2; 1.
 SMART; SMC0130; KR; 1.
 SMART; SMC0020; Tryp_Spc; 1.
 PROSITE; PS00022; EGF_1; 2.
 PROSITE; PS01186; EGF_2; 1.
 PROSITE; PS00026; EGF_3; 2.
 PROSITE; PS01253; FIBRONECTIN_1; 1.
 PROSITE; PS00023; FIBRONECTIN_2; 1.
 PROSITE; PS00021; KRINGLE_1; 1.
 PROSITE; PS00070; KRINGLE_2; 1.
 PROSITE; PS00240; TRYPSIN_DOM; 1.
 PROSITE; PS00134; TRYPSIN_HIS; 1.
 PROSITE; PS00135; TRYPSIN_SER; 1.
 Glycoproteins; Blood coagulation; Plasma; Kringle; Serine protease;
 Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
 NON_TER 1
 SIGNAL 1
 CHAIN 19 358 ALPHA-FACTOR X11A HEAVY CHAIN.
 CHAIN 359 603 ALPHA-FACTOR X11A LIGHT CHAIN.
 DOMAIN 46 87 FIBRONECTIN TYPE-II.
 DOMAIN 93 130 EGF-LIKE 1.
 DOMAIN 132 172 FIBRONECTIN TYPE-I.
 DOMAIN 173 209 EGF-LIKE 2.
 DOMAIN 216 294 KRINGLE.
 DOMAIN 312 342 PRO-RICH.
 DOMAIN 359 603 SERINE PROTEASE.
 ACT_SITE 398 398 CHARGE RELAY SYSTEM (BY SIMILARITY).
 ACT_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).
 ACT_SITE 551 551 CHARGE RELAY SYSTEM (BY SIMILARITY).
 DISULFID 97 109 BY SIMILARITY.
 DISULFID 103 118 BY SIMILARITY.
 DISULFID 120 129 BY SIMILARITY.
 DISULFID 134 162 BY SIMILARITY.
 DISULFID 160 169 BY SIMILARITY.
 DISULFID 177 188 BY SIMILARITY.
 DISULFID 182 197 BY SIMILARITY.
 DISULFID 199 208 BY SIMILARITY.
 DISULFID 216 294 BY SIMILARITY.
 DISULFID 237 276 BY SIMILARITY.

FT DISULFID 265 289 BY SIMILARITY.
 FT DISULFID 345 472 BY SIMILARITY.
 FT DISULFID 383 399 BY SIMILARITY.
 FT DISULFID 391 461 BY SIMILARITY.
 FT DISULFID 422 425 BY SIMILARITY.
 FT DISULFID 488 557 BY SIMILARITY.
 FT DISULFID 520 536 BY SIMILARITY.
 FT DISULFID 547 578 BY SIMILARITY.
 FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 603 AA; 66795 MW; 48DC6B946FB9ED59 CRC64;
 Query Match 31.0%; Score 700.5; DB 1; Length 603;
 Best Local Similarity 36.7%; Pred. No. 6.8e-49;
 Matches 160; Conservative 60; Mismatches 147; Indels 69; Gaps 13;
 QY 13 CLNGSTCVSNKYFNHWCNCPKFGGQHCEIDSKYCEGNGHYRKGASTDTMGPCPL 72
 DB 182 CLNGRCLE---VEGHLCDPCMGYTGPFCDLDTTASCYEGRGVSYRGMARTTVSGAKCQ 238
 QY 73 PWSATVLQQTVAHRSD-ALQLGLGKHNYCRNPNRRRPMCVYQVGLKPLVQECMVHDC 131
 DB 239 RWAS-----EATYNNMTAEQALRRGLGHTFCRNPNDTRPWCFFVMGNRLSWEYCDLAQC 294
 QY 132 -----ADGKLKFO-----CGQKTLRPRF 149
 DB 295 QYPPQATATPHDFEHPKLPSSRLSILOPQTQNALNELPETSSLLCGQR-LRKRL 353
 QY 150 ---KIIGGETTIENQWFAALYRRHGGSVTVVCGSLISPCWVISATHCFIDYPKKE 205
 DB 354 SLSRIVGLVALPGAHFYIALY----WGS--NFCSSLIAPCWLTAAHCLQNRPAPE 407
 QY 206 DYIVYLGSRNSNTQGMKFEVENILHKDYSADTLAHNDIALLLKI-RKSGECAQPS 264
 DB 408 ELKVVLGQDRNQSCHEQCOTLAHVSRLHEAFSPS--SYLNDLALLRLQKSDGSCAQLS 465
 QY 265 RTIOTICLPSMYNDPQFG--TSCETITGKENSIDYLYPEOLKMTVVKLISHRECQPHY 322
 DB 466 PYQIVTCLPSGAPPSESETTCCVAGHGFEGAEYSSFLQEAQVFLISERCSSDEV 525
 QY 323 YGSEVTTKMLCAADPWKTDSCQSGSGGLVLC---SLQGRMTLTGIVSGRGCAKDKRPG 379
 DB 526 HGDAFLSGVLGAGFLEGGTDACQSGGGLVCEDEAAEHRLLRGLVSGSGCGDRNFKG 585
 QY 380 VYTRYSHPLPWIRSH 395
 DB 586 VYTDVASYLTWIKHT 601
 RESULT 19
 FA12_HUMAN STANDARD; PRT; 615 AA.
 ID AC P00748; P78339;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
 DE (HAP).
 GN F12.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=68007593; PubMed=2888762;
 RA Cool D.E., McGilivray R.T.A.;
 RT "Characterization of the human blood coagulation factor XII gene.
 RT Intron/exon gene organization and analysis of the 5'-flanking
 RT region.";
 RL J. Biol. Chem. 262:13662-13673(1987).
 RN [2]

RF SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605.
R2 Rieder M.J., Arnel T.Z., Carrington D.P., Ozuna M., Kulidnek S.A.,
R3 Rajkumar N., Toth E.J., Yi O., Nickerson D.A.;
R4 Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
R5 [3]
R6 SEQUENCE OF 4-615 FROM N.A.
R7 MEDLINE=86176794; PubMed=3754331;
R8 Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
R9 Cortese R.;
R10 "cDNA sequence coding for human coagulation factor XII (Hageman).";
R11 Nucleic Acids. Res. 14:3146-3146(1986).
R12 [4]
R13 SEQUENCE OF 14-615 FROM N.A.
R14 MEDLINE=86033830; PubMed=3877053;
R15 Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
R16 McGillivray R.T.A.;
R17 "Characterization of human blood coagulation factor XII cDNA.
R18 Prediction of the primary structure of factor XII and the tertiary
R19 structure of beta-factor XIIa.";
R20 J. Biol. Chem. 260:13666-13676(1985).
R21 [5]
R22 SEQUENCE OF 146-615 FROM N.A.
R23 MEDLINE=86216049; PubMed=3011063;
R24 Que B.G., Davie E.W.;
R25 "Characterization of a cDNA coding for human factor XII (Hageman
R26 factor).";
R27 Biochemistry 25:1525-1528(1986).
R28 [6]
R29 SEQUENCE OF 20-379.
R30 MEDLINE=85182674; PubMed=3886654;
R31 McMullen B.A., Fujikawa K.;
R32 "Amino acid sequence of the heavy chain of human alpha-factor XIIa
R33 (activated Hageman factor).";
R34 J. Biol. Chem. 260:5328-5341(1985).
R35 [7]
R36 SEQUENCE OF 561-615 FROM N.A.
R37 TISSUE=Blood;
R38 MEDLINE=96133302; PubMed=8528215;
R39 Schloesser M., Hoferbert S., Bartz U., Lutze G., Lammle B., Engel W.;
R40 "The novel acceptor splice site mutation 11396(G-->A) in the factor
R41 XII gene causes a truncated transcript in cross-reacting material
R42 negative patients.";
R43 Hum. Mol. Genet. 4:1235-1237(1995).
R44 [9]
R45 CARBOHYDRATE-LINKAGE SITE THR-109.
R46 MEDLINE=92184750; PubMed=1544894;
R47 Harris R.J., Ling V.T., Spellman M.W.;
R48 "O-linked fucose is present in the first epidermal growth factor
R49 domain of factor XII but not protein C.";
R50 J. Biol. Chem. 267:5102-5107(1992).
R51 [10]
R52 VARIANT WASHINGTON D.C. SER-590.
R53 MEDLINE=90046788; PubMed=2510163;
R54 Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B.,
R55 Saito H.;
R56 "Coagulation factor XII (Hageman factor) Washington D.C.: inactive
R57 factor XIIa results from Cys-571-->Ser substitution.";
R58 Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).
R59 [11]
R60 VARIANT LOCARNO PRO-372.
R61 MEDLINE=94325559; PubMed=8049433;
R62 Hovanga J.K., Schaller J., Stricker H., Wullemijn W.A., Furian M.,
R63 Laemmle B.;
R64 "Coagulation factor XII Locarno: the functional defect is caused by
R65 the amino acid substitution Arg-353-->Pro leading to loss of a
R66 kallikrein cleavage site.";
R67 Blood 84:1173-1181(1994).

[12]
RP VARIANT TENRI CYS-53.
RX MEDLINE=99290785; PubMed=10361128;
RA Kondo S., Tokunaga F., Kawano S., Oono Y., Kumagai S., Koide T.;
RT "Factor XII Tenri, a novel cross-reacting material negative factor XII
RT deficiency, occurs through a proteasome-mediated degradation.";
RL Blood 93:4300-4308(1999).
CC -!- FUNCTION: Factor XII is a serum glycoprotein that participates in
CC the initiation of blood coagulation, fibrinolysis, and the
CC generation of bradykinin and angiotensin.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla.
CC -!- PTM: O- AND N-GLYCOSYLATED.
CC -!- DISEASE: Defects in F12 do not cause any clinical symptoms. The
CC sole effect is that whole-blood clotting time is prolonged. There
CC is a complex bound to an anionic surface. Prekallikrein forms a
CC complex bound to an anionic surface. Prekallikrein is cleaved by
CC factor XII to form kallikrein, which then cleaves factor XII first
CC to alpha-factor XIIa and then to beta-factor XIIa. Alpha-factor
CC XIIa activates factor XI to factor Xla.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M31315; AAA70225.1; -.
CC EMBL; AF538691; AAM97932.1; -.
CC EMBL; M11723; AA51986.1; -.
CC EMBL; M17466; AAB59490.1; -.
CC EMBL; M17464; AAB59490.1; JOINED.
CC EMBL; M17465; AAB59490.1; JOINED.
CC EMBL; M13147; AAA70224.1; -.
CC EMBL; U71274; AAB51203.1; -.
CC PIR; A29411; KEHU12.
CC HSSP; P00763; LDPO.
CC MEROPS; S01.211; -.
CC Genes; HGNC:3530; F12.
CC XIM; 234000; -.
CC GO; GO:0003805; F: blood coagulation factor XI activity; TAS.
CC GO; GO:0003806; F: blood coagulation factor XII activity; TAS.
CC GO; GO:0008236; F: serine-type peptidase activity; TAS.
CC GO; GO:0007596; P: blood coagulation; TAS.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR000083; Fibrinctnl.
CC InterPro; IPR000562; FN_Type_II.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1a.
CC Pfam; PF000008; EGF; 2.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00040; fn2; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00013; FNTYPEII.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000995; FN_Type_II; 1.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00059; FN2; 1.
CC SMART; SM00130; KR; 1.

```

ID FA12_BOVIN STANDARD; PRT; 593 AA.
AC P98140;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAP) (Fragment).
DE F12.
GN Bos taurus (Bovine).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
OX NCBI
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=liver;
RC MEDLINE=94242782; PubMed=8186251;
RX Shiba Y., Samba U., Okabe H., Kanbara T., Yamamoto T.;
RT "Primary structure of bovine Hageman factor (blood coagulation factor
RT XII): comparison with human and guinea pig molecules.";
RL Biochim. Biophys. Acta 1206:63-70(1994).
RN [2]
RP SEQUENCE OF 10-21; 350-364 AND 525-550.
RX MEDLINE=77182112; PubMed=861210;
RT Fujikawa K., Walsh A.K., Davie W.E.;
RT "Isolation and characterization of bovine factor XII (Hageman
RT factor)";
RL Biochemistry 16:2270-2278(1977).
CC -1- FUNCTION: Factor XII is a serum glycoprotein that participates in
CC the initiation of blood coagulation, fibrinolysis, and the
CC generation of bradykinin and angiotensin.
CC -1- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-ile bonds in factor
CC VII to form factor VIII and factor XI to form factor Xla.
CC -1- PPM: O- AND N-GLYCOSYLATED (BY SIMILARITY).
CC -1- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a
CC complex bound to an anionic surface. Prekallikrein is cleaved by
CC factor XII to form kallikrein, which then cleaves factor XII first
CC to alpha-factor Xlla and then to beta-factor Xlla. Alpha-factor
CC Xlla activates factor XI to factor Xla. Bovine factor XII is
CC cleaved only to alpha-factor Xlla as it lacks the trypsin/
CC kallikrein cleavage site.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S70164; BAB30804.2; --
DR PIR; S45281; S45281.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.211; --
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrctn1.
DR InterPro; IPR000562; FN Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF000008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.

```

PFAM; PF00051; K1INGRE; 1.
PFAM; PF00089; TRYPSIN; 1.
PRINTS; PR00722; CHYMOTRYPSIN.


```

very Match      22.5%; Score 508.5; DB 1; Length 790;
est Local Similarity 36.2%; Pred. No. 2.6e-33;
atches 130; Conservative 45; Mismatches 141; Indels 43; Gaps 14;

( 45 DKSCTCYEONGHFVAGKASTDTMGRCPLPNWSATVLQQTVYHAHR-----SDALQLGLGKHN 100
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
I 456 DLSDCMGNGKRYGKGAATVAGVPCQEWAA-----QEPHRHSIFTPETNPRAGLEK-N 509

( 101 YCRNPD-NRRRPWCVTQVGLKPLVQECMWHDCADGKKFCQGQKTLRPR---FKIIGEF 156
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
I 510 YCRNPDGDDNGPWCYT-TNPQKLFYCDVPCQVTS--SFDGCKPKVEPKCCPARVVGCV 566

( 157 TTIEQNPFAPATYYRHRGGSVTVYCGGSLSPCNVISATCFIDYPKKEDYTWLGRSL 216
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
I 567 SIPLSWPQWISIRYRYRG---HFCGGTLLSPENVLTAKHCLKSEKSSPSYKVLGAHEE 622

( 217 NSNTOGEMKPEVENIILHKVYSADTLAHNDIALLKIRSKEGRCAPFSRTIOTICLPSMY 276
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
623 YHLGEGVQVIDSKLFF--KEPS-----EADIALKLSPP-----AVITDKVIPACLPT-- 668

```


Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
"Angiostatin: a novel angiogenesis inhibitor that mediates the
suppression of metastases by a Lewis lung carcinoma.";
Cell 79:315-328(1994).
C-!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
a proteolytic factor in a variety of other processes including
embryonic development, tissue remodeling, tumor invasion, and
inflammation; in ovulation it weakens the walls of the Graafian
follicles. It activates the urokinase-type plasminogen activator,
collagenases and several complement zymogens, such as C1 and C5.
It cleaves fibrin, fibronectin, thrombospondin, laminin and von
Willebrand factor.
C-!- FUNCTION: Angiostatin is an angiogenesis inhibitor that blocks
neovascularization and growth of experimental primary and
metastatic tumors in vivo.
C-!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
higher selectivity than trypsin. Converts fibrin into soluble
products.
C-!- ENZYME REGULATION: Converted into plasmin by plasminogen
activators, both plasminogen and its activator being bound to
fibrin. Cannot be activated with streptokinase.
C-!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
immediately after dissociation from the clot.
C-!- MISCELLANEOUS: In the presence of the inhibitor, the activation
involves only cleavage after Arg-581, resulting in 2 chains held
together by 2 disulfide bonds. Without the inhibitor, the
activation involves also removal of the activation peptide.
C-!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
C-!- SIMILARITY: Contains 5 kringle domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; J04766; AAA50168.1; --
EMBL; AF481053; AAM22156.1; --
EMBL; BC014773; AAM14773.1; --
EMBL; BC057186; AAM57186.1; --
EMBL; AV134430; AAM15805.1; --
PIR; A38514; PLMS.
HSP; P00747; IPMK.
MEROPS; S01.233; --
MGD; GGI:97620; Plg.
GO; GO:0016506; F:apoptosis activator activity; IDA.
GO; GO:0006915; P:apoptosis; IDA.
InterPro; IPR009003; Cys Ser trypsin.
InterPro; IPR000001; Kringle.
InterPro; IPR003014; PAN.
InterPro; IPR003609; Pan_app.
InterPro; IPR001254; Peptidase_S1.
InterPro; IPR001314; Peptidase_S1A.
InterPro; IPR003966; Peptidase_S1A_pr.
Pfam; PF000051; kringle; 5.
Pfam; PF00024; PAN; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; KRINGLE.
PRINTS; PR01505; PROTHROMBIN.
ProDom; PD000395; Kringle; 4.
SMART; SM00130; KR; 5.
SMART; SM00473; PAN AP; 1.
SMART; SM00020; TRYD_SPC; 1.
PROSITE; PS00021; KRINGLE_1; 5.
PROSITE; PS00070; KRINGLE_2; 5.
PROSITE; PS0240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
K: Tissue remodeling; Blood coagulation; Plasma; Glycoprotein; Fibrinolysis;
Hydrolase; Serine protease; Kringle; Zymogen; Repeat;

[illegible]

345 QDGGGLVCSLQGRWTLTGIVSGRCALKDKRGVYTRVSHPLPHI 391
759 QDGGGLVCFVCKDKYILQVTLGSLGRCARPNKPGVYVRSRFDWI 805

ULT 23
N MACMU STANDARD; PRT; 810 AA.
P12545;
01-OCT-1989 (Rel. 12, Created)
01-OCT-1989 (Rel. 12, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Plasminogen precursor (EC 3.4.21.7).
PLG.
Macaca mulatta (Rhesus macaque).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Theria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecoidea; Macaca.
NCBI_TaxID=9544;
[1]
SEQUENCE FROM N.A.
MEDLINE=89174660; PubMed=2925643;
Tomlinson J.E., McLean J.W., Lawn R.M.;
"Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
synthesis.";
J. Biol. Chem. 264:5957-5965(1989).
-!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
a proteolytic factor in a variety of other processes including
embryonic development, tissue remodeling, tumor invasion, and
inflammation; in ovulation it weakens the walls of the Graafian
follicle. It activates the urokinase-type plasminogen activator,
collagenases and several complement zymogens, such as C1 and C5.
It cleaves fibrin, fibronectin, thrombospondin, laminin and von
Willebrand factor.
-!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-[Xaa]-Arg-[Xaa];
higher selectivity than trypsin. Converts fibrin into soluble
products.
-!- ENZYME REGULATION: Converted into plasmin by plasminogen
activators, both plasminogen and its activator being bound to
fibrin. Activated with catalytic amounts of streptokinase.
-!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
immediately after dissociation from the clot.
-!- MISCELLANEOUS: In the presence of the inhibitor, the activation
involves only cleavage after Arg-580, resulting in 2 chains held
together by 2 disulfide bonds. Without the inhibitor, the
activation involves also removal of the activation peptide.
-!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
-!- SIMILARITY: Contains 5 kringle domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

EMBL; J04697; AAA36901.1; -
PIR; B32869; B30848.
HSSP; P00747; 1PMK.
MEROPS; S01.233; -
InterPro; IPR009003; Cys Ser trypsin.
InterPro; IPR000001; Kringle.
InterPro; IPR003014; PAN.
InterPro; IPR003609; Pan_app.
InterPro; IPR001254; Peptidase S1.
InterPro; IPR001314; Peptidase_S1A.
InterPro; IPR003966; Peptidase_S1A_pr.
Pfam; PF00051; kringle; 5.
Pfam; PF00024; PAN; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00158; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR PRODOM; PD000395; Kringle; 5.
DR SMART; SMO0130; KR; 4.
DR SMART; SMO0473; PAN AP; 1.
DR SMART; SMO0020; TYD SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 5.
DR PROSITE; PS00070; KRINGLE 2; 5.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW Signal.
FT SIGNAL 1 19
FT CHAIN 20 810 PLASMINOGEN.
FT CHAIN 20 580 PLASMIN HEAVY CHAIN A.
FT PEPTIDE 20 96 ACTIVATION PEPTIDE.
FT CHAIN 97 580 PLASMIN SHORT FORM OF CHAIN A.
FT CHAIN 581 810 PLASMIN LIGHT CHAIN B.
FT DOMAIN 103 181 KRINGLE 1.
FT DOMAIN 184 262 KRINGLE 2.
FT DOMAIN 275 352 KRINGLE 3.
FT DOMAIN 377 454 KRINGLE 4.
FT DOMAIN 481 560 KRINGLE 5.
FT DOMAIN 581 810 SERINE PROTEASE.
FT ACT_SITE 622 622 CHARGE RELAY SYSTEM.
FT ACT_SITE 665 665 CHARGE RELAY SYSTEM.
FT ACT_SITE 760 760 CHARGE RELAY SYSTEM.
FT BINDING 136 136 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 158 158 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 172 172 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 432 432 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 445 445 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 134 134 FIBRIN.
FT BINDING 136 136 FIBRIN.
FT BINDING 49 73 BY SIMILARITY.
FT DISULFID 53 61 BY SIMILARITY.
FT DISULFID 103 181 BY SIMILARITY.
FT DISULFID 124 164 BY SIMILARITY.
FT DISULFID 152 176 BY SIMILARITY.
FT DISULFID 185 262 BY SIMILARITY.
FT DISULFID 188 316 BY SIMILARITY.
FT DISULFID 206 245 BY SIMILARITY.
FT DISULFID 234 257 BY SIMILARITY.
FT DISULFID 275 352 BY SIMILARITY.
FT DISULFID 296 335 BY SIMILARITY.
FT DISULFID 324 347 BY SIMILARITY.
FT DISULFID 377 454 BY SIMILARITY.
FT DISULFID 398 437 BY SIMILARITY.
FT DISULFID 426 449 BY SIMILARITY.
FT DISULFID 481 560 BY SIMILARITY.
FT DISULFID 502 543 BY SIMILARITY.
FT DISULFID 531 555 BY SIMILARITY.
FT DISULFID 567 685 BY SIMILARITY.
FT DISULFID 577 585 BY SIMILARITY.
FT DISULFID 607 623 BY SIMILARITY.
FT DISULFID 699 766 BY SIMILARITY.
FT DISULFID 729 745 BY SIMILARITY.
FT DISULFID 756 784 BY SIMILARITY.
FT CARBOHYD 365 365 O-LINKED (GALNAc...) (BY SIMILARITY).
SQ SEQUENCE 810 AA; 90255 MW; A75E1C51A1A0F24A CRC64;

Query Match 22.2%; Score 501; DB 1; Length 810;
Best Local Similarity 34.4%; Pred No. 1,le-32;
Matches 140; Conservative 46; Mismatches 165; Indels 56; Gaps 16;

QY 10 NCDCUNGTCVSNKYFSNIHWCNCPKFGQ-----HCEIDKSKTCYEGNGH 56
Db 428 NPDADKGPWCFTTDPDSVRWEYCNLKKCSGTGSAAPPVPAQLPDAETPSDEDCMFGNGK 487
QY 57 FYRGKASTDTWGRCLPWSNATVLOQTVHAHR-----SDALQGLGKHNYCNPED-NRRRP 111

488 GYRKGKATVTGTGTCQEWAA-----QEPHSHRIFTPETNPRAGLEK-NYCRNPDGVDGGP 541
112 WCYVQVGLKPLVQSCVHDCADGKLFQCGQKTLRPR---FKIIGGEFTTIENQPFAAI 168
542 WCYT-TNPRKLFYCDVPQCAAS--SFQCKPQVEPKPCGRVVGCVAYPHSWPW--QI 596
169 YRRHGGSTVYVCGSLISPCWISATHCTIDVPKEDYIVYIGRSLNSNTGEMKFEV 228
597 SLRTRLG--NHFCGGTILSPWVLTAAHCKSKSRPSFYKVLGAHR-----EVHLEPHV 649
229 ENLILHKDYSADTLAHENDTALKIRSKGRCQPSRTIQTICLPSMYNDPQS-----GTS 284
650 QEIVSRQMFSEPARA---DIALLKSSP-----AIITKVIPACLPSS-----PNVIVADPTE 698
285 CEITGFGKENSIDYLPQOLKMTVVKLISHRECOQPHYGYSEVTTKMLCAADPQWKDSC 344
699 CFITGWTGETQGT--YGAGLLKEARLPVIENKVCNRYEFLNGTVTKTTELCAHLAGGTDSC 756
345 QDSGGPLVCLSGRMVLTIGVSGRCALKDKRGVTVRSHLPWI 391
757 QDSGGPLVCFEKKYILQGVTSWGLGCLARPKNRGVYVRSRVFTWI 803

LT 24
PI HUMAN
II PLAN HUMAN STANDARD; PRT; 810 AA.
AC P00747;
D7 21-JUL-1986 (Rel. 01, Created)
D9 01-MAR-1989 (Rel. 10, Last sequence update)
D1 15-MAR-2004 (Rel. 43, Last annotation update)
D1 Plasmminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
G6 PLG.
O6 Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
R1 SEQUENCE FROM N.A., AND VARIANT ASN-472.
R2 MEDLINE=90202879; PubMed=2318848;
R3 Petersen T.E., Marten M.R., Ichinose A., Davis E.W.;
R4 "Characterization of the gene for human plasminogen, a key proenzyme
R5 in the fibrinolytic system.";
R6 J. Biol. Chem. 265:6104-6111(1990).
R7 [2]
R8 SEQUENCE FROM N.A.
R9 MEDLINE=87162490; PubMed=3030813;
R10 Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
R11 "Molecular cloning and characterization of a full-length cDNA clone
R12 for human plasminogen.";
R13 FEBS Lett. 213:254-260(1987).
R14 [3]
R15 SEQUENCE FROM N.A., AND VARIANTS LYS-57; GLN-133; HIS-261; TRP-408;
R16 ASN-472; VAL-494 AND TRP-523.
R17 Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
R18 Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
R19 Submitted (DEC-2002) to the ENBL/GenBank/DBJ databases.
R20 [4]
R21 SEQUENCE OF 20-810, AND VARIANT ASN-472.
R22 Sottrup-Jensen L., Petersen T.E., Magnusson S.;
R23 Submitted (JUL-1977) to the PIR data bank.
R24 [5]
R25 MEDLINE=85023311; PubMed=6148961;
R26 Malinowski D.P., Sadler J.E., Davis E.W.;
R27 "Characterization of a complementary deoxyribonucleic acid coding for
R28 human and bovine plasminogen.";
R29 Biochemistry 23:4243-4250(1984).
R30 [6]
R31 SEQUENCE OF 20-100.
R32 MEDLINE=75093329; PubMed=122932;
R33 Wiman B., Wallen P.;
R34 "Structural relationship between 'glutamic acid' and 'lysine' forms
R35 of human plasminogen and their interaction with the NH2-terminal

RT activation peptide as studied by affinity chromatography.";
RL Eur. J. Biochem. 50:489-494(1975).
RP [7]
R2 SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810, AND VARIANT ASN-472.
R3 Sottrup-Jensen L., Claess H., Zajdel M., Petersen T.E., Magnusson S.;
R4 (in) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);
R5 Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
R6 Raven Press, New York (1978).
R7 [8]
R8 SEQUENCE OF 483-604.
R9 MEDLINE=76043692; PubMed=1268663;
R10 Wiman B., Wallen P.;
R11 "Amino-acid sequence of the cyanogen-bromide fragment from human
R12 plasminogen that forms the linkage between the plasmin chains.";
R13 Eur. J. Biochem. 58:539-547(1975).
R14 [9]
R15 SEQUENCE OF 581-810.
R16 MEDLINE=77225245; PubMed=142009;
R17 Wiman B.;
R18 "Primary structure of the B-chain of human plasmin.";
R19 Eur. J. Biochem. 76:129-137(1977).
R20 [10]
R21 ACTIVE SITE.
R22 MEDLINE=73149248; PubMed=4694729;
R23 Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
R24 "The primary structure of human plasminogen. II. The histidine loop
R25 of human plasmin: light (B) chain active center histidine sequence.";
R26 J. Biol. Chem. 248:1631-1633(1973).
R27 [11]
R28 ACTIVE SITE.
R29 MEDLINE=69234739; PubMed=4240117;
R30 Groeskopf W.R., Summaria L., Robbins K.C.;
R31 "Studies on the active center of human plasmin. Partial amino acid
R32 sequence of a peptide containing the active center serine residue.";
R33 J. Biol. Chem. 244:3590-3597(1969).
R34 [12]
R35 OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
R36 MEDLINE=82213905; PubMed=6919539;
R37 Trexler M., Vali Z., Pathy L.;
R38 "Structure of the omega-aminocarboxylic acid-binding sites of human
R39 plasminogen. Arginine 70 and aspartic acid 56 are essential for
R40 binding of ligand by kringle 4.";
R41 J. Biol. Chem. 257:7401-7406(1982).
R42 [13]
R43 FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
R44 MEDLINE=85054794; PubMed=6094526;
R45 Vali Z., Pathy L.;
R46 "The fibrin-binding site of human plasminogen. Arginines 32 and 34
R47 are essential for fibrin affinity of the kringle 1 domain.";
R48 J. Biol. Chem. 259:13690-13694(1984).
R49 [14]
R50 PHOSPHORYLATION SITE SER-597.
R51 MEDLINE=97345939; PubMed=9201958;
R52 Wang H., Prorok M., Bretthauer R.K., Castellino F.J.;
R53 "Serine-578 is a major phosphorylation locus in human plasma
R54 plasminogen.";
R55 Biochemistry 36:8100-8106(1997).
R56 [15]
R57 CARBOHYDRATE-LINKAGE SITES.
R58 MEDLINE=88185329; PubMed=3356193;
R59 Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
R60 Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
R61 "The N- and O-linked carbohydrate chains of human, bovine and porcine
R62 plasminogen. Species specificity in relation to sialylation and
R63 fucosylation patterns.";
R64 Eur. J. Biochem. 173:57-63(1988).
R65 [16]
R66 CARBOHYDRATE-LINKAGE SITE SER-268.
R67 MEDLINE=97207306; PubMed=9054441;
R68 Pirie-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
R69 Pizzo S.V.;
R70 "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of
R71 human plasminogen 2.";

J. Biol. Chem. 272:7408-7411(1997).
[17]
CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
MEDLINE=95042728; PubMed=7525077;
O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
"Angiostatin: a novel angiogenesis inhibitor that mediates the
suppression of metastases by a Lewis lung carcinoma.";
Cell 79:315-328(1994).
[18]
CHARACTERIZATION OF ANGIOSTATIN.
MEDLINE=97239710; PubMed=9102221;
Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,
Lapcevich R., Nacy C.A.;
"A recombinant human angiotatin protein inhibits experimental primary
and metastatic cancer.";
Cancer Res. 57:1329-1334(1997).
[19]
X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
MEDLINE=92031502; PubMed=1657148;
Mulichak A.M., Tulinsky A., Ravichandran K.G.;
"Crystal and molecular structure of human plasminogen kringle 4
refined at 1.9-A resolution.";
Biochemistry 30:10576-10588(1991).
[20]
X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
MEDLINE=92031503; PubMed=1657149;
Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
"The refined structure of the epsilon-aminocaproic acid complex of
human plasminogen kringle 4.";
Biochemistry 30:10589-10594(1991).
[21]
X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
Sec B., Yamano A., Whitlow M., Teeter M.M.;
"Structure of human plasminogen kringle 4 at 1.68 Angstrom and 277 K.
A possible structural role of disordered residues.";
Acta Crystallogr. D 53:169-178(1997).
[22]
X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.
MEDLINE=96180681; PubMed=8611560;
Mathews I.I., Vandernhoff-Hanover P., Castellino F.J., Tulinsky A.;
"Crystal structures of the recombinant kringle 1 domain of human
plasminogen in complexes with the ligands epsilon-aminocaproic acid
and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
Biochemistry 35:2567-2576(1996).
[23]
X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
MEDLINE=99198034; PubMed=9521645;
Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
Castellino F.J.;
"Structure and ligand binding determinants of the recombinant kringle
5 domain of human plasminogen.";
Biochemistry 37:3258-3271(1998).
[24]
STRUCTURE BY NMR OF 96-184.
MEDLINE=94237157; PubMed=8181475;
Rejante M.R., Llinas M.;
"1H-NMR assignments and secondary structure of human plasminogen
kringle 1.";
Eur. J. Biochem. 221:927-937(1994).
[25]
STRUCTURE BY NMR OF 96-184.
MEDLINE=94237158; PubMed=8181476;
Rejante M.R., Llinas M.;
"Solution structure of the epsilon-aminohexanoic acid complex of
human plasminogen kringle 1.";
Eur. J. Biochem. 221:939-949(1994).
[26]
STRUCTURE BY NMR OF 183-354.
MEDLINE=96194156; PubMed=8652577;
Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
Rickli E.E.;
"Recombinant gene expression and 1H NMR characteristics of the

RT kringle (2 + 3) supermodule: spectroscopic/functional individuality
RT of plasminogen kringle domains.";
RL Biochemistry 35:2357-2364(1996).
RN [27]
RP STRUCTURE BY NMR OF 374-461.
RX MEDLINE=90219023; PubMed=2157850;
Query Match 22.1%; Score 499; DB 1; Length 810;
Best Local Similarity 33.7%; Pred. No. 1.6e-32;
Matches 137; Conservative 44; Mismatches 170; Indels 56; Gaps 15;
QY 10 NCDCLNGTGVSNKYFSNIHWCNCPKXFGGO-----HCEIDSKTCYEGNGH 56
Db 428 NPDADKGMCTTDPSEWEYCNLKCSGTSEASVAVPPVLLPDVETPSEDCWFGNGK 487
QY 57 FYRGKASDTMGRPCLPWNSATVLQOTYHAHR----SDALQLGLGKNYCNPD-NRRRP 111
Db 488 GYRGKATVTGTTCQDWA-----QEPHRSITPTETNPRAGLEK-NYCRNPQDVGCP 541
QY 112 WYVQVGLKPLVQECMVHDCADGLKFCQGTQKTLRPR---FKIIGBFTTIENQWFAAI 168
Db 542 WCYT-TNPKLYDYCDVQPCA--APSFDCGKQVEPKKPCGRVVGCGVAHPSWQVSL 598
QY 169 YRRHGGSVTVYCGSLSPCWVISATHCFIDYPKEDYIVYLGSRSLNNTQGMKEEV 228
Db 599 --RTRFG--MHFCGTLISPEWVITAHCLKSPSPSYKVILGAHOEVNLEPHVQEEV 654
QY 229 ENLIHKDYSADTLAHNDIALKIRSKEGRCAPSRITQICLPSMTNDPQF----GTS 284
Db 655 SRLFLEPT-----RKDIALKLSSP---AVITDKVIPACLPS----PNVVADRTE 698
QY 285 CEITGFGKENSVDYLYPEQLKTVVVKLISHRECOQPHYGVSEVITMCAADPOWKDSC 344
Db 699 CFITGWGTQGT--FGAGLLKEAQLPVIEKNRYEFLNGRVQSTELCAGHLAGGTDSC 756
QY 345 QQDSGGPLVCSLQGRMTLTGTVSGRGCAKDKPGVYTRVSHFLPWI 391
Db 757 QQDSGGPLVCFEKOKYILQGVTSWGLGCAKPNKPGVYVRVSEFVTWI 803
RESULT 25
PLMN CANFA
ID PLMN CANFA STANDARD; PRT; 333 AA.
AC P80009;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragment).
GN PLG.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RC TISSUE=Plasma;
RX MEDLINE=90175323; PubMed=2626424;
Schaller J., Straub C., Kaempfer U., Rickli E.E.;
"Complete amino acid sequence of canine miniplasminogen.";
Protein Seq. Data Anal. 2:445-450(1989).
RL -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
a proteolytic factor in a variety of other processes including
embryonic development, tissue remodeling, tumor invasion, and
inflammation; in ovulation it weakens the walls of the Graafian
follicle. It activates the urokinase-type plasminogen activator,
collagenases and several complement zymogens, such as C1 and C5.
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von
CC Willembrand factor.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen
CC activators, both plasminogen and its activator being bound to

Db 215 VVADRLCYITGWGEFGQT- -YGAGLLKEAQLPVIEKNCVRYYEYLNGRVKSTELCAGNL 272

QY 338 QWKTDSCQDGSGLPVCSLQGRWTLTGIVSWRGCCALKDKPGYVTRVSHPLPW 391

Db 273 AGGTDSCQDGSGLPVCSLQGRWTLTGIVSWRGCCALKDKPGYVTRVSHPLPW 326

RESULT 26

ID	PLMN_SHEEP	STANDARD;	PRT;	343 AA.
AC	P81286;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DE	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Plasminogen (EC 3.4.21.7) (Fragment)			
GN	PLG.			
OS	Ovis aries (Sheep)			
OC	Eukaryota; Metazoa;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Caprinae; Ovis.			
ON	NCBI_TaxID=9940;			
RX	[1]			
RP	SEQUENCE.			
RX	MEDLINE=93149995; PubMed=1492092;			
RA	Schaller J., Straub C., Kampfer U., Rickli E.B.;			
RT	"Complete amino acid sequence of ovine miniplasminogen.";			
RL	Protein Seq. Data Anal. 5:21-25(1992).			
CC	-!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as a proteolytic factor in a variety of other processes including embryonic development, tissue remodeling, tumor invasion, and inflammation; in ovulation it weakens the walls of the Graafian follicle. It activates the urokinase-type plasminogen activator, collagenases and several complement zymogens, such as C1 and C5. It cleaves fibrin, fibronectin, thrombospondin, laminin and von Willebrand factor.			
CC	-!- CATALYTIC ACTIVITY: Preferential cleavage: Lys- -Xaa > Arg- -Xaa; higher selectivity than trypsin. Converts fibrin into soluble products.			
CC	-!- ENZYME REGULATION: Converted into plasmin by plasminogen activators, both plasminogen and its activator being bound to fibrin. CANNOT be activated with streptokinase.			
CC	-!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin immediately after dissociation from the clot.			
CC	-!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.			
CC	-!- SIMILARITY: Contains at least 2 kringle domains.			
DR	PIR; B61545; B61545.			
DR	HSSP; P00747; SHP.			
DR	MEROPS; S01.233; -.			
DR	InterPro; IPR009003; Cys Ser trypsin.			
DR	InterPro; IPR000001; Kringle_			
DR	InterPro; IPR001254; Peptidase S1.			
DR	InterPro; IPR001314; Peptidase_S1A.			
DR	InterPro; IPR003966; Peptidase_S1A_pr.			
DR	Pfam; PF00051; kringle; 1.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; P00722; CHYMOTRYPSIN.			
DR	PRINTS; P00018; KRINGLE.			
DR	PRINTS; P01505; PROTHROMBIN.			
DR	ProDom; PD000395; Kringle; 1.			
DR	SMART; SM00130; KR; 1.			
DR	SMART; SM00020; TRY_SPC; 1.			
DR	PROSITE; PS00021; KRINGLE_1; 1.			
DR	PROSITE; PS00070; KRINGLE_2; 1.			
DR	PROSITE; PS02040; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;			
KW	Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.			
FT	NON_TER 1			
FT	DOMAIN <1 140 HEAVY CHAIN A.			
FT	DOMAIN 141 >343 LIGHT CHAIN A.			
FT	DOMAIN <1 17 KRINGLE 4.			

FT DOMAIN 110 188 KRINGLE 1.
 FT DOMAIN 192 269 KRINGLE 2.
 FT DOMAIN 282 359 KRINGLE 3.
 FT DOMAIN 384 461 KRINGLE 4.
 FT DOMAIN 485 564 KRINGLE 5.
 FT DOMAIN 584 812 SERINE PROTEASE.
 FT CARBOHYD 315 315 /FTID=CAR 000014.
 FT CARBOHYD 365 365 /FTID=CAR 000015.
 FT ACT_SITE 624 624 CHARGE RELAY SYSTEM.
 FT ACT_SITE 667 667 CHARGE RELAY SYSTEM.
 FT ACT_SITE 762 762 CHARGE RELAY SYSTEM.
 FT CONFLICT 335 335 N -> D (IN REF. 2).
 FT CONFLICT 516 516 Q -> H (IN REF. 2).
 FT CONFLICT 555 555 P -> L (IN REF. 2).
 FT CONFLICT 744 744 T -> R (IN REF. 3).
 SC SEQUENCE 812 AA; 51216 MW; 386A6A91E220946 CRC64;
 ary Match 21.5%; Score 484.5; DB 1; Length 812;
 at Local Similarity 35.4%; Pred. No. 2.3e-31;
 ches 123; Conservative 42; Mismatches 151; Indels 31; Gaps 11;
 QY 50 CYEGNGHYRGKASPTDTRGCLPNSATVLQOTVHAHRSDALQGLGKHNCRNPD-NR 108
 DL 485 CMIGTKSYRGKATTVAGVPCQEWAEQPHQSIFTPETNP-OSGL-ERNYCRNPDGV 542
 QY 109 RRPWCYVQGLKPLVQECWHDGADGKLFQCGQKTLRPR---PKIIGGEFTTIENQPF 165
 DL 543 NGPWCYTNPRKPF-DYCDVQPC---ESSFCGPKPKVEPKKSGRIYGGCVKSPHSPWQ 598
 QY 166 AAIYRRHGGVYVCGSLSPCWISATHCDFIDYFKEDYIVYLGSRSLNSVTQEMK 225
 DL 599 VSLRSSR-----HFCGGTLISPRKVLTAACHLDNLIALSYKVLGAHNEKREQVQE 653
 QY 226 FEVENLILHKQYSADTLAHHNDIALKIRSKEGRCAPFSRTIQTICLPSMYNDPQFQTS 285
 DL 654 IPVSLFREPSEA-----DIALKL-----SRPAITKEVIFACLPNPMVAARTEC 701
 QY 286 EITGKGNSTNLYPE-OLKMTVVKLISHRECQOPHYGSEVTKMLCAADPQWKTDC 344
 DL 702 YITGNETQGT---FGEGLLKEAHLPIENKVCNREYLDGRVKPTELCAHLLGGTDC 758
 QY 345 QDSGGPLVCSLQGRMTLTGVSWRGCAKDKPGVYTRVSHFLPW 391
 DL 759 QDSGGPLVCFEKDKVILQGVTSWGLGCAKPNKGVYRVSPYVPI 805
 RE LT 28
 AI HUMAN
 IL APOA HUMAN STANDARD; PRT; 4548 AA.
 AC P085T9;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).
 GN LPA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OY NCBI_TaxID=9606;
 RA [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88039109; PubMed=3670400;
 RA McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,
 RA Pless G.M., Scanu A.M., Lawn R.M.;
 RJ "cDNA sequence of human apolipoprotein(a) is homologous to
 RV plasminogen.";
 RI Nature 330:132-137 (1987).
 RN [2]
 RN SERINE PROTEASE ACTIVITY.
 RJ MEDLINE=90076123; PubMed=2531657;
 RJ Salonen E.-M., Jaahainen M., Zardi L., Vaheri A., Ehnholm C.;

RT "Lipoprotein(a) binds to fibronectin and has serine proteinase
 activity capable of cleaving it.";
 RL EMBO J. 8:4035-4040(1989).
 RN [3]
 RP REVIEW.
 RX MEDLINE=90049223; PubMed=2530631;
 RA Utermann G.;
 RT "The mysteries of lipoprotein(a).";
 RL Science 246:904-910(1989).
 RN [4]
 RP CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.
 RX MEDLINE=21303595; PubMed=11294842;
 RA Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;
 RT "Structural elucidation of the N- and O-glycans of human
 apolipoprotein(a): role of o-glycans in conferring protease
 resistance.";
 RL J. Biol. Chem. 276:22200-22208 (2001).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.
 RX MEDLINE=96217891; PubMed=8642595;
 RA Mikol V., Lograsso P.V., Boettcher B.R.;
 RT "Crystal structures of apolipoprotein(a) kringle IV37 free and
 complexed with 6-aminocaproic acid and with p-aminomethylbenzoic
 acid: existence of novel and expected binding modes.";
 RL J. Mol. Biol. 256:751-761 (1996).
 RN [6]
 RP VARIANT ARG-4193.
 RX MEDLINE=95002201; PubMed=7918682;
 RA Scanu A.M., Pfaffinger D., Lee J.C., Hinman J.;
 RT "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37
 associated with a lysine binding defect in Lp(a).";
 RL Biochim. Biophys. Acta 1227:41-45 (1994).
 CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
 (Lp(a)). It has serine proteinase activity and is able of
 autolysis. Inhibits tissue-type plasminogen activator 1.
 CC Lp(a) may be a ligand for megalin/Gp 330.
 CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
 decorin.
 CC -!- PTM: N- and O-glycosylated. The N-glycans are complex biantennary
 structures present in either a mono- or disialylated state. The
 O-glycans are mostly (80%) represented by the monosialylated core
 type I structure, NeuNacalpha2-3Galbeta1-3GalNAc, with smaller
 amounts of disialylated and non-sialylated O-glycans also
 detected.
 CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its
 naturally occurring proteolytic fragments are correlated with
 atherosclerosis. Homology with plasminogen kringle IV and V is
 thought to underlie the atherogenicity of the protein, because the
 fragments are competing with plasminogen for fibrin(ogen) binding.
 CC -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
 leading to the formation of the so called mini-Lp(a). Apo(a)
 fragments accumulate in atherosclerotic lesions, where they may
 promote thrombogenesis. O-glycosylation may limit the extent of
 proteolytic fragmentation.
 CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
 CC -!- SIMILARITY: Contains 38 kringle domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X06290; CAA29618.1; -.
 DR PIR; S00657; S00657.
 DR PDB; 1I71; 13-JUN-01.
 DR PDB; 1JFN; 28-JUN-02.
 DR PDB; 1KIV; 18-MAY-99.
 DR PDB; 3KIV; 18-MAY-99.
 DR PDB; 4KIV; 18-MAY-99.
 DR MEROPS; S01.226; -.

Genew; HGNC:6667; LPA.
MIM; 152200; ..
GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
GO; GO:0008015; P:circulation; TAS.
GO; GO:0009405; P:patogenesis; TAS.
InterPro; IPR009003; Cys_Ser_trypsin.
InterPro; IPR000001; Kringle.
InterPro; IPR001254; Peptidase_S1.
InterPro; IPR001314; Peptidase_S1A.
Pfam; PF00051; Kringle; 38.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; KRINGLE.
ProDom; PD000395; Kringle; 38.
SMART; SM00130; KR; 38.
SMART; SM00020; Tryp_Spc; 1.
PROSITE; PS00021; KRINGLE 1; 38.
PROSITE; PS50070; KRINGLE 2; 38.
PROSITE; PS50240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
HydroLase; Serine protease; Lipid transport; Plasma; Glycoprotein;
Kringle; Repeat; Atherosclerosis; Signal; Polymorphism; 3D-structure.
SIGNAL 1 19
F CHAIN 20 4548 APOLIPOPROTEIN(A).
F DOMAIN 20 130 KRINGLE TYPE IV, 1.
F DOMAIN 131 244 KRINGLE TYPE IV, 2.
F DOMAIN 245 358 KRINGLE TYPE IV, 3.
F DOMAIN 359 472 KRINGLE TYPE IV, 4.
F DOMAIN 473 586 KRINGLE TYPE IV, 5.
F DOMAIN 587 700 KRINGLE TYPE IV, 6.
F DOMAIN 701 815 KRINGLE TYPE IV, 7.
F DOMAIN 815 928 KRINGLE TYPE IV, 8.
F DOMAIN 929 1042 KRINGLE TYPE IV, 9.
F DOMAIN 1043 1156 KRINGLE TYPE IV, 10.
F DOMAIN 1157 1270 KRINGLE TYPE IV, 11.
F DOMAIN 1271 1384 KRINGLE TYPE IV, 12.
F DOMAIN 1385 1498 KRINGLE TYPE IV, 13.
F DOMAIN 1499 1612 KRINGLE TYPE IV, 14.
F DOMAIN 1613 1726 KRINGLE TYPE IV, 15.
F DOMAIN 1727 1840 KRINGLE TYPE IV, 16.
F DOMAIN 1841 1954 KRINGLE TYPE IV, 17.
F DOMAIN 1955 2068 KRINGLE TYPE IV, 18.
F DOMAIN 2069 2182 KRINGLE TYPE IV, 19.
F DOMAIN 2183 2296 KRINGLE TYPE IV, 20.
F DOMAIN 2297 2410 KRINGLE TYPE IV, 21.
F DOMAIN 2411 2524 KRINGLE TYPE IV, 22.
F DOMAIN 2525 2638 KRINGLE TYPE IV, 23.
F DOMAIN 2639 2752 KRINGLE TYPE IV, 24.
F DOMAIN 2753 2866 KRINGLE TYPE IV, 25.
F DOMAIN 2867 2980 KRINGLE TYPE IV, 26.
F DOMAIN 2981 3094 KRINGLE TYPE IV, 27.
F DOMAIN 3095 3208 KRINGLE TYPE IV, 28.
F DOMAIN 3209 3322 KRINGLE TYPE IV, 29.
F DOMAIN 3323 3436 KRINGLE TYPE IV, 30.
F DOMAIN 3437 3550 KRINGLE TYPE IV, 31.
F DOMAIN 3551 3664 KRINGLE TYPE IV, 32.
F DOMAIN 3665 3778 KRINGLE TYPE IV, 33.
F DOMAIN 3779 3892 KRINGLE TYPE IV, 34.
F DOMAIN 3893 3998 KRINGLE TYPE IV, 35.
F DOMAIN 3999 4112 KRINGLE TYPE IV, 36.
F DOMAIN 4113 4226 KRINGLE TYPE IV, 37.
F DOMAIN 4227 4340 KRINGLE TYPE V.
F DOMAIN 4341 4454 SERINE PROTEASE.
F ACT SITE 4369 CHARGE RELAY SYSTEM.
F ACT SITE 4412 CHARGE RELAY SYSTEM.
F ACT SITE 4498 CHARGE RELAY SYSTEM.
F VARIANT 4193 W -> R (LOSS OF LYSINE-SEPHAROSE BINDING)
F SEQUENCE 4548 AA; 501313 MW; 96921BE96A465C5F CRC64;
F very Match 21.2%; Score 478; DB 1; Length 4548;

Best Local Similarity 36.0%; Pred. No. 5.Se-30;
Matches 129; Conservative 36; Mismatches 133; Indels 60; Gaps 17;
QY 50 CYEGNGHYRGKASTDTHGEPCLPWNATVLTQTYHAHRSALQLGLGK-----HNYCRN 104
DB 4228 CMFGNGKGYRGKATVTGTPCQEW-----AAQEPHRSHTFIPGNKAGLEKNTCRN 4280
QY 105 PD-NRRPWCYVQGLKPLVQECVHDCADGKLKFOGQKTLRPR---FKIIGSEFTTIE 160
DB 4281 PDGDSINGPWCYT-MNPRKLFYCDIPLCASS--SFDGCKPQVEPKPGSGIVGCVVAHPH 4337
QY 161 NOPWFAALYRHRHGGSVTVYVGGSLSPCWVTSATHCFDYPKKEDYVVLGRSR---LN 217
DB 4338 SWPQVSL--RTRPGK--HFCGGTLSPWVLTAAHCKLKSSRPSSFKVILGAHQEVNLE 4393
QY 218 SNTQEMKFEVENLILHKDYSADTLAHHNDIALIKIRSEKRCQAQPSRTIQTICLPSMYN 277
DB 4394 SHVQ--EIEVSRLEPTQA-----DIALKL---SRPAVITDKVMPACLPS--- 4435
QY 278 DPQF-----GTSCETITGFKENSTDYLPOLKMTVVKLVSHRECCQHHYGVSEVTTKMLC 333
DB 4436 -PDYMTVARTECYITGWGEVQGT--FTGLLKAQLLVIEVNCN--HY-----KYIC 4483
QY 334 AADPQWKTDCSQGDSGGPLVCSLQGRMTLTGIVSWRGKCALDKPKGVYTVRVSHPLPMI 391
DB 4484 AEHLARGTDCSQGDSGGPLVCFEKKYILQGVTSWGLGCARPKNKPGVYARVSRFVTWI 4541
RESULT 29
ID APOA MACMU STANDARD; PRT; 1420 AA.
AC P14417;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (Fragment).
GN LPA.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_taxid=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89174660; PubMed=2925643;
RA Tomlinson J.E., McLean J.W., Lawn R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
synthesis.";
RL J. Biol. Chem. 264:5957-5965(1989).
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
(Lp(a)). It has serine proteinase activity and is able of
autoproteolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
decorin (By similarity).
CC -!- PTM: N- and O-glycosylated (By similarity).
CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its
naturally occurring proteolytic fragments are correlated with
atherosclerosis. Homology with plasminogen kringle IV and V is
thought to underlie the atherogenicity of the protein, because the
fragments are competing with plasminogen for fibrinogen binding.
CC -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
leading to the formation of the so called mini-Lp(a). Apo(a)
fragments accumulate in atherosclerotic lesions, where they may
promote thrombogenesis. O-glycosylation may limit the extent of
proteolytic fragmentation (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -!- SIMILARITY: Contains at least 10 kringle domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; J04635; AAA36833.1; ..
DE DE PIR; A32869; A32869.
DE DE HSSP; P00747; 2PK4.
DE DE MEROPS; S01.226; -.
DE DE InterPro; IPRO09003; Cys Ser trypsin.
DE DE InterPro; IPRO00001; Kringle-
DE DE InterPro; IPRO001254; Peptidase_S1.
DE DE InterPro; IPRO01314; Peptidase_S1A.
DE DE Pfam; PF00051; kringle; 11.
DE DE Pfam; PF00089; trypsin; 1.
DE DE PRINTS; PRO0722; CHYMOTRYPSIN.
DE DE PRINTS; PRO0018; KRINGLE.
DE DE ProDom; PD000395; Kringle; 10.
DE DE SMART; SM00130; KR; 10.
DE DE SMART; SM00020; TRYD_SPC; 1.
DE DE PROSITE; PS00021; KRINGLE_1; 10.
DE DE PROSITE; PS00070; KRINGLE_2; 10.
DE DE PROSITE; PS00240; TRYPsin DOM; 1.
DE DE PROSITE; PS00134; TRYPsin HIS; FALSE NEG.
DE DE PROSITE; PS00135; TRYPsin SER; FALSE NEG.
DE DE KxHydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
DE DE Kringle; Repeat; Atherosclerosis.
FT NON PER 1 1
FT DOMAIN 49 127 KRINGLE 1.
FT DOMAIN 163 241 KRINGLE 2.
FT DOMAIN 277 355 KRINGLE 3.
FT DOMAIN 391 469 KRINGLE 4.
FT DOMAIN 505 583 KRINGLE 5.
FT DOMAIN 619 697 KRINGLE 6.
FT DOMAIN 725 803 KRINGLE 7.
FT DOMAIN 839 917 KRINGLE 8.
FT DOMAIN 953 1031 KRINGLE 9.
FT DOMAIN 1067 1145 KRINGLE 10.
FT DOMAIN 1191 1420 SERINE PROTEASE.
SC SEQUENCE 1420 AA; 158367 MW; BEI02949E03CB0E CRC64;
ery Match 21.1%; Score 477; DB 1; Length 1420;
st Local Similarity 34.0%; Pred. No. 1.7e-30;
tches 129; Conservative 43; Mismatches 137; Indels 70; Gaps 15;
Qy 50 CYGNGHGYRKASTDNGRCLPWNSATVQL-QTYAHRSDALQLGLGHKNTCRPNPDN 107
De 1068 CYHGNGSGYRGTFSTTVTGRTCQSWSSTPHQHKTPTENHPNDLTM-----NYCRNPDA 1122
Qy 108 RRRRWCVQVLKLVSQ--CWHDCAQ-----GKLKFCQG 142
De 1123 DTGNWCFI--WDSEVRREYNLFECSDTEGVTPPTVPVPSLEAPSEQASSFDCGK 1179
Qy 143 KTLAPR---FKIIIGEFTTIENQWFAAIYPRHRGGSVYYVCGSLISPCWVISATHCFI 199
De 1180 PQVFPEKKCFGSIAGCVAAHPSPWQVSL--RTRPGK--HFCGGTLLISPEWVLTAAACGLE 1235
Qy 200 DYPKKEDIVYVLRGR---LNSNTQGENMKFEVENILHKDYASDTLAHNDDIALLKIRSK 256
De 1236 TFSRPSFKVILGAHQEVNLSSHVQ---EIEVSRFLFEPIGA-----DJALLKL---- 1281
Qy 257 EGRCAQPRTQTICLPWNDDPF-----GTSCETITGFGKENSTDIYPEOLKMTVVKLI 312
De 1282 -SRFAIITDKVIPACLPS-----PNYVITAWTECYITGWGETQGT--FGAGLLKEAQLHVI 1334
Qy 313 SHRECQQPHYYSGEVTTTKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRCG 372
De 1335 ENTVCNHVEFLNGRVKSTELCAGHLAGGTDRCCQDNGDGPPVCFDKDKILAGITSWGFGC 1394
Qy 373 ALKOKPGVYTRVSHFLPMI 391
De 1395 ACNKPGEVYVRVSSFTWI 1413

RESULT 30	PLMN HORSE	STANDARD;	PRT;	338 AA.
ID	PLMN HORSE	STANDARD;	PRT;	338 AA.
AC	P80010;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
GN	Plasminogen (EC 3.4.21.7) (Fragment).			
DN	PLG.			
OS	Equus caballus (Horse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.			
OX	NCB1_TaxID=9796;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=plasma;			
RC	MEDLINE=92052077; PubMed=1946332;			
RX	Schaller J., Straub C., Kaempfer U., Rickli E.E.;			
RA	"Complete amino acid sequence of equine miniplasminogen.";			
RL	Protein Seq. Data Anal. 4:69-74(1991).			
-1-	FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as a proteolytic factor in a variety of other processes including embryonic development, tissue remodeling, tumor invasion, and inflammation; in ovulation it weakens the walls of the Graafian follicle. It activates the urokinase-type plasminogen activator, collagenase and several complement zymogens, such as C1 and C5. It cleaves fibrin, fibronectin, thrombospondin, laminin and von Willebrand factor.			
-1-	CATALYTIC ACTIVITY: Preferential cleavage: Lys- -Xaa > Arg- -Xaa; higher selectivity than trypsin. Converts fibrin into soluble products.			
-1-	ENZYME REGULATION: Converted into plasmin by plasminogen activators, both plasminogen and its activator being bound to fibrin. Activated with catalytic amounts of streptokinase.			
-1-	MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin immediately after dissociation from the clot.			
-1-	SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily. PR; A61545; A61545.			
DR	HSP; P00747; SHPG.			
DR	MEROPS: S01.233; -.			
DR	InterPro: IPR009003; Cys_Ser_trypsin.			
DR	InterPro: IPR000001; Kringle_5.			
DR	InterPro: IPR001254; Peptidase_S1.			
DR	InterPro: IPR001314; Peptidase_S1A.			
DR	InterPro: IPR003966; Peptidase_S1A_pr.			
DR	Pfam: PF00051; kringle; 1.			
DR	Pfam: PF00089; trypsin; 1.			
DR	PRINTS: PR00722; CHYMOTRYPSIN.			
DR	PRINTS: PR00018; KRINGLE			
DR	PRINTS: PR03505; PROTHROMBIN.			
DR	ProDom: PD000395; Kringle; 1.			
DR	SMART: SM00130; KR; 1.			
DR	SMART: SM00020; Tryp_SPC; 1.			
DR	PROSITE: PS00021; KRINGLE 1; 1.			
DR	PROSITE: PSS0070; KRINGLE 2; 1.			
DR	PROSITE: PSS0240; TRYPSIN_DOM; 1.			
DR	PROSITE: PS00134; TRYPSIN_HIS; 1.			
DR	PROSITE: PS00135; TRYPSIN_SRR; 1.			
KW	Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;			
KW	Tissue remodeling; Blood coagulation; Kringle; zymogen.			
FT	NON TER 1 1			
FT	CHAIN <1 108			
FT	CHAIN 109 338			
FT	DOMAIN 9 88			
FT	DOMAIN 109 338			
FT	DISULFID 9 88			
FT	DISULFID 30 71			
FT	DISULFID 59 83			
FT	DISULFID 95 213			
FT	DISULFID 105 113			
FT	DISULFID 135 151			
FT	DISULFID 227 294			
FT	PLASMIN HEAVY CHAIN A.			
FT	PLASMIN LIGHT CHAIN B.			
FT	KRINGLE 5.			
FT	SERINE PROTEASE.			
FT	BY SIMILARITY.			
FT	BY SIMILARITY.			
FT	BY SIMILARITY.			
FT	INTERCHAIN (BY SIMILARITY).			
FT	INTERCHAIN (BY SIMILARITY).			
FT	BY SIMILARITY.			
FT	BY SIMILARITY.			

```
--!- TISSUE SPECIFICITY: Brian and Leydig cells of the testis.
--!- SIMILARITY: Belongs to peptidase family S1.
--!- SIMILARITY: Contains 1 kringle domain.
--!- SIMILARITY: Contains 4 SRCR domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL: AJ001531; CAA04816.1; --
EMBL: AF077298; RAD25919.1; --
HSP: P00783; IDPO.
Genew: HGNC:9477; PRSS12.
MIM: 606709; --
MEROPS: S01.237; --
GO: GO:0008236; F-serine-type peptidase activity; TAS.
InterPro: IPR009003; Cys Ser_trypsin.
InterPro: IPR000001; Kringle.
InterPro: IPR001254; Peptidase_S1.
InterPro: IPR001314; Peptidase_S1A.
InterPro: IPR001190; Srcr_receptor.
Pfam: PF00051; kringleg; 1.
Pfam: PF00530; SRCR; 4.
Pfam: PF00089; trypsin; 1.
PRINTS: PRO0722; CHYMOTRYPSIN.
PRINTS: PRO0258; SPERACTRCPTR.
ProDom: PD000395; Kringleg; 1.
SMART: SM00130; KR; 1.
SMART: SM00202; SR; 4.
SMART: SM00020; Tryp Spc; 1.
PROSITE: PS00021; KRINGLE_1; 1.
PROSITE: PS00070; KRINGLE_2; 1.
PROSITE: PS00420; SRCR_1; 3.
PROSITE: PS00287; SRCR_2; 4.
PROSITE: PS00240; TRYPSIN_DOM; 1.
PROSITE: PS00134; TRYPSIN_HIS; 1.
PROSITE: PS00135; TRYPSIN_SER; 1.
Hydrolase; Serine protease; Glycoprotein; Kringle; Repeat; Signal.
FT SIGNAL      1    20     POTENTIAL.
FT CHAIN       21   875     NEUROTIRYPSIN.
FT DOMAIN      23    92     PRO-RICH.
FT DOMAIN      93   165     KRINGLE.
FT DOMAIN     170   271     SRCR 1.
FT DOMAIN     280   381     SRCR 2.
FT DOMAIN     387   487     SRCR 3.
FT DOMAIN     500   601     SRCR 4.
FT DOMAIN     619   875     SERINE PROTEASE.
FT DOMAIN     619   875     ZMOGEN ACTIVATION REGION.
FT ACT_SITE    630   631     REACTIVE BOND (POTENTIAL) .
FT ACT_SITE    676   676     CHARGE RELAY SYSTEM.
FT ACT_SITE    726   726     CHARGE RELAY SYSTEM.
FT ACT_SITE    825   825     CHARGE RELAY SYSTEM.
FT DISULFID    619   750     POTENTIAL.
FT CARBOHYD     26    26     N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD    683   683     N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CONFLICT    663   663     A -> V (IN REF. 2) .
FT CONFLICT    701   701     E -> V (IN REF. 2) .
FT CONFLICT    839   841     VVV -> AAL (IN REF. 2) .
FT CONFLICT    875 AA; 97011 MW; B6SEC946DC208DC8 CRC64;
SEQUENCE
Query Match          19.8%; Score 447; DB 1; Length 875;
Best Local Similarity 29.7%; Pred. No. 2.6e-28;
Matches 128; Conservative 60; Mismatches 155; Indels 88; Gaps 17;
-----
QY 33 CPKTKGGQH-C-EIDKSNTCYEG-----NGHF--- 57
DB 465 CSRQRWGRHDCSHREDVSIACYPGEGHRLSLGFPVRLMDGENKKRGREVFINGOWGTI 524
QY 58 -----YRKASDTDM---GRPCLPWNSATVLQQTIVIAHRSDALQLGL 96
```

525 CDGWTDAVICRQYKGPARTMAVFGSGKPIHVDNVKCTGNSLADCIKQDI 584
 97 GKNNYCRNPDRRPPWYQVGLKPLVQECWHDGADG-----KLKFGCGQKTLRPFK- 150
 585 GRN-CHSEADAGVICY--FGKK-----ASGNSNKESLSVCGRLUHRQRK 630
 151 IIGSEFTIENQPFIAIYRRHRGGSVYVCGSLISPCWVGATGCHFDYPRK-EDYIV 209
 631 IIGKNSLRGGWPQVSLRLKSSHGDRLLCGATLLSSCWLVTAAHCFKRYGNSRYSYAV 690
 210 YLGRSLRNSNTQGMKEVENLILKDYSDATLAHNDIALIKRKEGRCAPSTIQ 269
 691 RVGYDHTLVPEEFEEIGVQIVIHREYRPRSDY--DIALVRLQPEEQCARFSSHVLP 748
 270 ICLFSMYNDPO-FGTSCEITGFGKENSTDYLYPEQLQMTVVKLISHRECCQPHYGVSEVT 328
 749 ACLPWRERPKTASNYITGNG---DTGRAYSETLQQAALPLPKFCEE--RYKGRFT 803
 329 TKMLCAAD--PQWNTDSCQDSGGLVCSLQGR-MTLTGIVSWGRGCGALCKDEGVTRVS 385
 804 GRMLCAGNLHKEKVDSCQDSGGLPCERPGESWVYVGTSGYCGVGVKDSFGVTRVS 863
 386 HFLPWIRSHTK 396
 864 AFVFWIKSVTK 874

IT 32
 MOUSE
 NEUTR MOUSE STANDARD; PRT; 761 AA.
 O0872;
 30-MAY-2000 (Rel. 39, Created)
 30-MAY-2000 (Rel. 39, Last sequence update)
 15-MAR-2004 (Rel. 43, Last annotation update)
 Neurotysin precursor (EC 3.4.21.-) (Motopsin) (Brain-specific serine
 protease 3) (BSSP-3).
 PRS12 OR BSSP3.
 Mus musculus (Mouse).
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 SEQUENCE FROM N.A.
 TISSUE=Brain;
 MEDLINE=97401523; PubMed=9245503;
 Geschwend T.P., Krueger S.R., Kozlov S.V., Wolfer D.P., Sonderegger P.,
 "Neurotysin, a novel multidomain serine protease expressed in the
 nervous system.",
 Mol. Cell. Neurosci. 9:207-219(1997).
 [2]
 SEQUENCE FROM N.A.
 MEDLINE=98008948; PubMed=9344839;
 Yamamura Y., Yamashiro K., Teurukua N., Nakazato H., Tsujimura A.,
 Yamaguchi N.,
 "Molecular cloning of a novel brain-specific serine protease with a
 kringle-like structure and three scavenger receptor cysteine-rich
 motifs.",
 Biochem. Biophys. Res. Commun. 239:386-392(1997).
 [3]
 SEQUENCE FROM N.A.
 STRAIN=C57BL/6J; TISSUE=Mammary gland;
 MEDLINE=22388257; PubMed=12477932;
 Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Udell T.B., Ioshizuka S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
 Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettman M., Madao A., Rodrigues S., Sanchez A.,
 Whiting M., Madao A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield V.S.N., Krzyzanski M.L., Skalska U., Smalhus D.E.,
 Smerich A., Schein J.E., Jones S.J.M., Maira M.A.,
 "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences",
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC - FUNCTION: Plays a role in neuronal plasticity and the proteolytic
 action may subserve structural reorganizations associated with
 learning and memory operations.
 CC - SUBCELLULAR LOCATION: Secreted.
 CC - TISSUE SPECIFICITY: Most abundant in cerebral cortex, hippocampus
 and amygdala.
 CC - SIMILARITY: Belongs to peptidase family S1.
 CC - SIMILARITY: Contains 1 kringle domain.
 CC - SIMILARITY: Contains 3 SRCR domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 or send an email to license@isb-sib.ch).
 EMBL; Y13192; CAA73646.1; -
 EMBL; D89871; BAA23966.1; -
 EMBL; BC031429; AAB31429.1; -
 HSSP; P00763; 1DPO.
 DR MEROPS: S01.237; -
 DR MGD: MGI:1100881; Press12.
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR InterPro: IPR001190; Srcr_receptor.
 DR Pfam: PF00051; Kringle; 1.
 DR Pfam: PF00530; SRCR; 3.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR ProDom: PD00256; SPERACTRCPTR.
 DR SMART: SM00130; KX; 1.
 DR SMART: SM00020; TRYPSIN_DOM; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS00420; SRCR_1; 3.
 DR PROSITE: PS00287; SRCR_2; 3.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR Hydrolase; Serine protease; Glycoprotein; Kringle; Repeat; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 761
 FT DOMAIN 85 157
 FT DOMAIN 166 267
 FT DOMAIN 273 373
 FT DOMAIN 386 487
 FT DOMAIN 505 761
 FT DOMAIN 505 516
 FT ACT_SITE 516 517
 FT ACT_SITE 562 562
 FT ACT_SITE 612 612
 FT ACT_SITE 711 711
 FT DISULFID 505 636
 FT CARBOHYD 93 93
 FT CARBOHYD 521 521
 FT CARBOHYD 569 569
 FT SEQUENCE 761 AA; 84118 MW; DF507B03712164E6 CRC64;

Query Match 19.2%; Score 434; DB 1; Length 761;
est Local Similarity 29.0%; Pred. No. 2.5e-27;
atches 126; Conservative 62; Mismatches 159; Indels 88; Gaps 16;
28 IHCNCPKPGGGQH-C--EIDKSKTCY-EGNGH-----56
346 VSPICQSRQWGRHDSHREDVGLTCYPSDGHRLSPGPIRLVDGENKKGRVEVFVNG 405
57 -----FYRGKASTDTW---GRFCLPWNSATVLQQTTHAHRSDA 91
406 QWGTICDDGWTDXAAVICRQLGKGPARTWAFEGEGKGPIMDNVXCTGNEKALADC 465
92 LQLGLGHKXNCRPNDRRRWCYVQ-----VGLKPLVQSCVHDCADGLKQCCQKTLR 146
466 VKQIDIGHN-CRSEDAVGICDYLEKASSGNKEM-----LSSGCGRLRLH 511
147 PRFK-IIGGEFTTIENQWPAALYRHRGGSVTVCGSLIPSCWVISATHCFIDY-PRK 204
512 RQKRIIGNNLSRGAWPQASLRSAHGDGRLGATLSSCWLTAAHCFKRYGNN 571
205 EDIVILGRRLNSNTGEMKFEVENILHKYSADTLAHNDIALKTRSEKRCQAQPS 264
572 RSYAVRVGDYHTLVPEEFEEIGVQOIVIHNRYPDRSDY--DIALVRLQGPGEQCARLS 629
265 RTIQTICLPSMYNDPO-FGTSCEITGFGKENSIDYLPKMTVVKLIASHRECCQPHY 323
630 THVLPACLPLWRPXPQTASCHITGWG---DTGRAYSTRLOQAAPLPLPKFCKE--RY 584
324 GSEVTRKMLCAADPQW-KTDSQCGSGGGLVCSLQGR-MTLGIVSWGRGCKALDKPGV 380
685 KGLFTGRMLCAGNLQEDNRVDSQCGSGGLMCKEKPDESVMVYGVTSWGYCGVXDTPGV 744
381 YTRVSHPLPWIRSH 395
745 YTRVPAFPVWIKSVT 759
ULT 33
N_ERIEU STANDARD; PRT; 810 AA.
Q29485;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Plasminogen precursor (EC 3.4.21.7).
PLG.
Erinaceus europaeus (Western European hedgehog).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceinae; Erinaceus.
[1] -
SEQUENCE FROM N.A.
TISSUE=Liver;
MEDLINE=96025778; PubMed=7592597;
Lawn R.M., Boommark N.W., Schwartz K., Lindahl G.E., Wade D.P.,
Byrne C.D., Fong K.J., Meer K.J., Patchy L.;
"The recurring evolution of lipoprotein(a). Insights from cloning of
hedgehog apolipoprotein(a).";
J. Biol. Chem. 270:24004-24009 (1995).
[2]
REVISIONS.
Lawn R.M.;
Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
a proteolytic factor in a variety of other processes including
embryonic development, tissue remodeling, tumor invasion, and
inflammation; in ovulation it weakens the walls of the Graafian
follicle. It activates the urokinase-type plasminogen activator,
collagenases and several complement zymogens, such as C1 and C5.
It cleaves fibrin, fibronectin, thrombospondin, laminin and von
Willebrand factor.
-!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;

higher selectivity than trypsin. Converts fibrin into soluble
products.
-!- ENZYME REGULATION: Converted into plasmin by plasminogen
activators, both plasminogen and its activator being bound to
fibrin. Cannot be activated with streptokinase.
-!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
immediately after dissociation from the clot.
-!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
-!- SIMILARITY: Contains 5 kringle domains.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
EMBL; U33171; AAC48717.1; --
PIR; I46260; I46260.
HSSP; P00747; 1PMK.
MEROPS; S01.233; --
DR InterPro; IPR009003; Cys_ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; Kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR PRODOM; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 5.
DR PROSITE; PS00070; KRINGLE 2; 5.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolyase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 810 PLASMINOGEN.
FT CHAIN 20 582 PLASMIN HEAVY CHAIN A (BY SIMILARITY).
FT CHAIN 583 810 PLASMIN LIGHT CHAIN B (BY SIMILARITY).
FT DOMAIN 583 810 SERINE PROTEASE.
FT DOMAIN 103 181 KRINGLE 1.
FT DOMAIN 185 262 KRINGLE 2.
FT DOMAIN 275 352 KRINGLE 3.
FT DOMAIN 379 456 KRINGLE 4.
FT DOMAIN 482 561 KRINGLE 5.
FT ACT_SITE 622 622 CHARGE RELAY SYSTEM.
FT ACT_SITE 665 665 CHARGE RELAY SYSTEM.
FT ACT_SITE 760 760 CHARGE RELAY SYSTEM.
FT CARBOHYD 339 339 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 810 AA; 90902 MW; 8E75780946017A16 CRC64;

Query Match 19.1%; Score 431; DB 1; Length 810;
Best Local Similarity 30.7%; Pred. No. 4.6e-27;
Matches 122; Conservative 49; Mismatches 157; Indels 70; Gaps 16;
QY 27 NIHW--CNCPKKFGQGHCEIDKS-----KTCVEGNHGFYRGKASTDTMGRPL 72
DB 445 SVRWFECNLKCKSGTEMATNSSPVQVSSASESSQDCIIDNGKGYRGTKATGTAGTPCQ 504
QY 73 PWSNATVLOQTHAH-----RSDALQLGLGHKHNCRAPD-NRRRPWCIVQVLKPL 122

505 AWA -----QEPHRSIFTPETNPRADL-----QENYCRNPGDANGPWCYT-TNPRKL 552

123 VQECWHDGADGKLFQCGQKTLRPRPKI---IGBETTIENOPFAAIYRRHRGGSVTY 179

553 FDYCDIPHCVSPS-SADCGPKVEPK-KCGRVGGCVVAHPHWPQVSLRRFQ-----H 605

180 VCGSLSPCWISATCFIDPKEDYIYVLG---RSRLNSNTQ--GEMKFEVENLILH 234

606 FCGTLLSPWVVAACHLCKFSPNPAIKYKVLGHAQETRLERDVQIKGVTKMFL----- 560

235 KQVSADTLAHNDIALKIRSKGRCQAPRTQTICLPSMYNDPQFSCETITGSKEN 294

661 -----PYRADIALKLSP-----AIITDKQHPACLPSNTWVADRSLCYITGKERTK 708

295 STDVLYPEQKMTVVVKLISHRECOOPHYGSEVTTKMLCAADPQWKTDCQGSGLPLVC 354

709 GT--YGAGLLKEAQLPVENKVCNRQSFNGRVRSTELCAGHLAGGVDSQGSGLPLVC 766

355 SLOGRMTLTGVNGRCALKDKPGVYTRVSHLFWIR 392

767 FEKRYILQGVTSNGLCARLTRPGVYTRVSRVYVWLQ 804

UT 24

HUMAN

16-OCT-2001 (Rel. 40, Created)

16-OCT-2001 (Rel. 40, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

Airway trypsin-like protease precursor (EC 3.4.21.-).

HAT.

Homo sapiens (Human).

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;

SEQUENCE FROM N.A. PubMed=9565616;

MEDLINE=9823382; PubMed=9070615; Tauchihasi S., Ogawara M.,

Masuda K.-I., Yamaoka K., Takahashi M., Sano T.;

"Purification, characterization, and localization of a novel

trypsin-like protease found in the human airway";

Am. J. Respir. Cell Mol. Biol. 16:300-308(1997).

FUNCTION: May play some biological role in the host defense system

on the mucous membrane independently of or in cooperation with

other substances in airway mucous or bronchial secretions.

CATALYTIC ACTIVITY: Preferentially cleaves the C-terminal side of

arginine residues at the P1 position of certain peptides, cleaving

Boc-phe-Ser-Arg-4-methylcoumaryl-7-amide most efficiently and

having an optimum pH of 8.6 with this substrate.

ENZYME REGULATION: Strongly inhibited by diisopropyl

fluorophosphate, leupeptin, antipain, aprotinin, and soybean

trypsin inhibitor, but hardly inhibited by secretory leukocyte

protease inhibitor at 10 microM.

SUBUNIT: Monomer.

SUBCELLULAR LOCATION: Type II membrane protein. Activated by

cleavage and secreted.

TISSUE SPECIFICITY: Located in the cells of the submucosal serous

glands of the bronchi and trachea.

SIMILARITY: Belongs to peptidase family S1.

SIMILARITY: Contains 1 SEA domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL; AB002134; BAA28691.1; -.

DR HSSP; P00750; 1RTF.

DR MEROPS; S01.301; -.

DR MIM; 605349; -.

DR GO; GO:0005576; C:extracellular; TAS.

DR GO; GO:0005887; C:integral to plasma membrane; TAS.

DR GO; GO:0008233; F:peptidase activity; TAS.

DR GO; GO:0007585; P:respiratory gaseous exchange; TAS.

DR InterPro; IPR009003; Cys Ser trypsin.

DR InterPro; IPR001254; Peptidase S1.

DR InterPro; IPR001314; Peptidase_S1A.

DR InterPro; IPR000082; SEA_domain.

DR Pfam; PF01390; SEA; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00200; SEA; 1.

DR SMART; SM00020; Tryp_SPC; 1.

DR PROSITE; PS50024; SEA; 1.

DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_SER; 1.

DR PROSITE; PS00135; TRYPSIN_HIS; 1.

DR Hydrolase; Serine protease; Transmembrane; Signal-anchor; Zymogen;

KW Glycoprotein.

CHAIN 1 186 AIRWAY TRYPSIN-LIKE PROTEASE, NON-

CATALYTIC CHAIN.

FT CHAIN 187 418 AIRWAY TRYPSIN-LIKE PROTEASE, CATALYTIC

FT CHAIN

FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 21 41 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)

FT (POTENTIAL).

FT DOMAIN 42 418 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 44 164 SEA.

FT DOMAIN 187 417 SERINE PROTEASE.

FT ACT_SITE 227 227 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 368 368 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT DISULFID 173 292 INTERCHAIN (BY SIMILARITY).

FT DISULFID 212 228 BY SIMILARITY.

FT DISULFID 337 353 BY SIMILARITY.

FT DISULFID 364 393 BY SIMILARITY.

FT CARBOHYD 144 144 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 418 AA; 46263 MW; F4BC1DB020CFBD0 CRC64;

Query Match 18.9%; Score 425.5; DB 1; Length 418;

Best Local Similarity 34.1%; Pred. No. 66-27;

Matches 92; Conservative 53; Mismatches 102; Indels 23; Gaps 7;

QY 127 MVHDCADGKLFQCGQKTLRPRFKIIGGFTTIENQPMFAIYRRHRGGSVTVCGSLI 186

DB 169 LINECAGDPLTLTSEQ-----RIIGGTAEAGSWPMQVSLRNN-----AHHCGLI 217

QY 187 SFCWVISAPHCIDYFKPKEDYIYVLGRSLNSNTQGMKFEVENLILHKDYASDTLAHN 246

DB 218 NNWMLTAAHCFRNSNPRDWIATSG---ISTTPKLMRVRNLIHNNYKSAI--HEN 271

QY 247 DIALKIRSKGRCQAPRTQTICLPSMYNDPQFSCETITGSKENSTDVLYPEQLK 306

DB 272 DIALVRLNS-----VTFTKDHSVCLPAATQIPFGSTAYVTGWAQEVAGHTVPE-LRQ 326

QY 307 TVVKLISHRECOOPHYGSEVTTKMLCAADPQWKTDCQGSGLPLVCGLQGRM-TLGI 365

DB 327 GQVRIISNDVCNAPHSYNGAILSGMLCAGVPQGGVDACQGSGLPLVQEDSRRLFWIGI 386

QY 366 VSWGRGCAKDKPGVYTRVSHLFWIRSH 395

DB 387 VSWGDCGLPDRPGVYTRVSHLFWIRSH 416

DR	PROSITE; PS01209; LDLRA_1; 1.	AC	O35453; Q9C9W7;
DR	PROSITE; PS0068; LDLRA_2; 3.	DT	15-JUL-1998 (Rel. 36, Created)
DR	PROSITE; PS0240; TRYPSIN_DOM; 1.	DT	10-OCT-2003 (Rel. 42, Last sequence update)
DF	PROSITE; PS00134; TRYPSIN_HIS; 1.	DT	10-OCT-2003 (Rel. 42, Last annotation update)
DF	PROSITE; PS00135; TRYPSIN_SER; 1.	DE	Serine protease hepsin (EC 3.4.21.-).
KX	Glycolase; Serine protease; Repeat; Signal-anchor; Transmembrane;	GN	HPN
KX	Glycoprotein; Alternative splicing;	OS	Mus musculus (Mouse).
FT	DOMAIN 1 55	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
FT	TRANSMEM 56 76	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
FT		OX	NCBI_TaxID=10090;
FT		RN	[1] SEQUENCE FROM N.A. (ISOFORM 2).
FT		RP	SEQUENCE FROM N.A. (ISOFORM 2).
FT		RC	TISSUE=Liver;
FT		RX	MEDLINE=98058912; PubMed=9395459;
FT		RA	Vu T.-K.H., Liu R.W., Haakema C., Tonaasek J.J., Howard E.W.;
FT		RA	"Identification and cloning of the membrane-associated serine
FT		RT	protease, hepsin, from mouse preimplantation embryos.";
FT		RL	Eur. J. Biol. Chem. 272:31315-31320(1997).
FT		RN	[2]
FT		RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
FT		RX	MEDLINE=99339944; PubMed=10411637;
FT		RA	Kawamura S., Kurachi S., Dayashiki Y., Kurachi K.;
FT		RT	"Complete nucleotide sequence, origin of isoform and functional
FT		RT	characterization of the mouse hepsin gene.";
FT		RL	Eur. J. Biochem. 262:755-764(1999).
FT		RN	[3]
FT		RP	SEQUENCE FROM N.A. (ISOFORM 1).
FT		RC	STRAIN=C57BL/6J; TISSUE=Kidney;
FT		RX	MEDLINE=21085660; PubMed=11217851;
FT		RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
FT		RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
FT		RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
FT		RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
FT		RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
FT		RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
FT		RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
FT		RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Wasio T.,
FT		RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
FT		RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
FT		RA	Brownstein M.J., But C., Fletcher C., Fujita M., Gariboldi M.,
FT		RA	Gustingich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
FT		RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
FT		RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
FT		RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
FT		RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
FT		RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
FT		RA	Hayashizaki Y.;
FT		RL	"Functional annotation of a full-length mouse cDNA collection.";
FT		RL	Nature 409:685-690(2001).
CC	-!- FUNCTION: Plays an essential role in cell growth and maintenance	CC	-!- of cell morphology.
CC	-!- SUBCELLULAR LOCATION: Type II membrane protein.	CC	-!- ALTERNATIVE PRODUCTS:
CC	Event=Alternative splicing; Named isoforms=2;	CC	Name=1; Synonyms=1a;
CC	Isoid=O35453-1; Sequence=Displayed;	CC	Note=Minor isoform;
CC	Name=2; Synonyms=2a;	CC	Isoid=O35453-2; Sequence=VSP_007232;
CC	Note=Major isoform;	CC	-!- SIMILARITY: Belongs to peptidase family S1.
CC	-!- CAUTION: Ref.3 sequence differs from that shown due to	CC	frameshifts in positions 155, 191 and 233.
CC	-----	CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
CC	the European Bioinformatics Institute. There are no restrictions on its	CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).	CC	-----
CC	-----	CC	-----
DR	EMBL; AF030065; AAB84221.1; -.		

QY	2	NELHQPNSDCLNGGTCVSNKYFSNIHWC-----NCPKFGGHCEDKSKTCVGGNGHF	57
DL	453	NQSDPCGEGFLCSVNGLCVPA-----CDGVKDCNGIDRNCVCRAFFQCKEDSTCI	504
QY	58	YRGKASTDTMGRP-CLPWSATVLIQ-----TYHAHRSDALQLGLGKHNKYNCRPNRR	109
DL	505	SLPKVCD---GQPDCLNGSDEQCQGVPCGTFTFQCEDRSCKK-----PNPQCDG	553
QY	110	RPWCYVQVGLKPLVQECWHDPCADGKLFQCGQKTLRPFKTIIGGEFTTIENOPWFAAY	169
DL	554	RP-----DCRGSDEHCDGLOGPSSRIVGGAVSSGEPWQASLQ	595
QY	170	RRHGGSVYVCGSLISPCWVISATHCFIDYPKKED-----YIVYLGSRSLNSNTOG	222
DL	596	VRGR-----HTCGALLADRWTAAHCF-----QEDSMASITVLWTFLGKWNQSRWPG	645
QY	223	EMKPEVENLIHKDYSADTLAHHNDIALLKRSKEGCAQPSRTIQICLPMSYNDPQFG	282
DL	646	EVSPKSRLLHPHDEED---SHDYDVALLQDHPVR-----SAARVPCVLPASHFFEPG	699
QY	283	TSCEITGFG---KENSTDYLYPEQLKMTVVKLIISRECCQPHYTGSEVTTMLCAADPQWK	340
DL	700	LHCWITWGALREGGP---ISNALQKVDVLIPODLCSEVRY---QVTPMLCAGYRKKG	754
QY	341	TDSCGSGGGPLVC-SLOGRMTLGIYSGWGCALKDKPGYTVRSHFLNIR	392
DL	755	KDACQSGSGGPLVCKAUSGRWFLAGLSWGLGCGRPNYFGVYTRITGVISWIQ	807
RI	LT 36		
HF	_MOUSE		
II	_HEPS_MOUSE	STANDARD; PRT; 436 AA.	

EMBL; AK002694; BAB22289.2; ALT_FRAME.

HSP; P00763; IDPO.

MEROFS; S01.224; -.

MGI; 1196620; Hpn.

InterPro; IPR009003; Cys Ser trypsin.

InterPro; IPR001254; Peptidase S1.

InterPro; IPR001314; Peptidase S1A.

InterPro; IPR001190; Src_rceptor.

Pfam; PF00089; trypsin; 1.

PRINTS; P00722; CHYMOTRYPSIN.

SMART; SM00202; SR; 1.

SMART; SM00202; TRYP_SPC; 1.

PROSITE; PS02440; TRYPsin DOM; 1.

PROSITE; PS00134; TRYPsin HIS; 1.

PROSITE; PS00135; TRYPsin SER; 1.

Hydrolase; Serine protease; Transmembrane; Signal-anchor;

Alternative splicing.

CHAIN 1 181

CHAIN 182 436

DOMAIN 21 36

TRANSMEM 37 63

DOMAIN 64 436

DOMAIN 182 436

ACT_SITE 222 222

ACT_SITE 276 276

ACT_SITE 372 372

ACT_SITE 372 372

DISULFID 172 296

DISULFID 207 296

DISULFID 341 357

DISULFID 368 400

CARBOHYD 131 131

VARSPLIC 25 44

CONFLICT 85 85

CONFLICT 204 204

CONFLICT 214 214

CONFLICT 228 229

CONFLICT 264 264

CONFLICT 281 281

SEQUENCE 436 AA; 46787 MW; 4A0993148C620BD0 CRC64;

very Match 18.5%; Score 417; DB 1; Length 436;

est Local Similarity 34.8%; Pred. No. 3e-26;

atches 111; Conservative 59; Mismatches 98; Indels 52; Gaps 18;

112 WCYVQVGLKPLVQECM----VHDCADGKLF---CQCKTLRPRFKIIGEEFTTIENQP 163

137 FCVDEGGL-PLAQRLLDVISVDCRGFELTATCDGGRKL-FVDRIVGQSDSLGRWP 194

164 WFAALYRHRGGSVTVYCGSLISPCWVISATHCFIDYPKEDVI-----VVLGRSLNSN 219

195 WQVSL--RYDG---THLCGSLSGDWLTAHCF---PERNVLRSRWFVAG--AVART 244

220 TQGMKPEVENLILH-----KDYSDATLAHNDIALKIRSKGRCACQPSRTTQTICLP 273

245 SPHAVQLGVQAVIYHGGVLPDPDPTID--ENSNDIALVHLSSS-----LPLTEYIQVCLP 298

274 SMYNDPQPTGTCHEITGFKENSTVLYPEQ---LMTVVVKLISHRECOQPHYGEVTK 330

299 AAGQALVDGKVTWTGNGNTQ----FYGQAVLQEARVPIISNEVCNSPDPYGNQIKPK 354

331 MLCRADPQWKTDCOGDSGGLVC--SLQG--RWLTGIVSGRCALKDKPGVTVRVSH 386

355 MFCAGYPGGIDACOGDSGGPPVCBDSISGTSRWLGLVSWGTGCLALRPGVTVTKTD 414

387 FLPW----IRSHKTEENGL 401

415 PREWIFKAIKTHS-EASGM 432

RESULT 37

TMS6 MOUSE

ID TMS6 MOUSE

AC Q9DBI0;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Transmembrane protease, serine 6 (EC 3.4.21.-) (Matrptase-2).

GN TMPRSS6.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.

RC STRAIN=C57BL/6J;

RC MEDLINE=22755759; PubMed=12744720;

RA Hooper J.D.; Campagnolo L., Goodarzi G., Truong T.N., Stuhlmann H.,

RA Quigley J.P.;

RA "Mouse matritase-2: identification, characterization and comparative

RT mRNA expression analysis with mouse hepsin in adult and embryonic

RT tissues.";

RL Biochem. J. 373:689-702(2003).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Liver;

RC MEDLINE=22354683; PubMed=12466851;

RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,

RA Nakaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,

RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,

RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,

RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,

RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,

RA Grimmond S., Gustinchik S., Hirokawa N., Jackson I.J., Jarvis E.D.,

RA Kanai A., Kawai H., Kawasaki Y., Kedzierski R.M., King B.L.,

RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,

RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,

RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,

RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,

RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,

RA Sandelin A., Schneider C., Semple C.A., Setou M., Shmida K.,

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,

RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,

RA Wilming L.G., Wyszewski-Boris A., Yanagisawa M., Yang I., Yang L.,

RA Hirozane-Kishikawa T., Kohno H., Nakamura M., Sakazume N., Sato K.,

RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,

RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,

RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,

RA Vasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

RA Birney E., Hayashizaki Y.;

RA "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RC MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Narusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Boek S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield I.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [4]
 RE REVIEW
 RX MEDLINE=22658120; PubMed=12784999;
 RA Netzel-Arnett S., Hooper J.D., Szabo R., Madison E.L., Quigley J.P.,
 RA Bugge T.H., Antalis T.M.;
 RT "Membrane anchored serine proteases: a rapidly expanding group of cell
 surface proteolytic enzymes with potential roles in cancer.";
 Cancer Metastasis Rev. 22:237-258(2003).
 CC -!- FUNCTION: May play a specialized role in matrix remodeling
 processes in liver (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
 CC -!- TISSUE SPECIFICITY: Expressed at highest levels in liver, kidney
 and uterus.
 CC -!- DEVELOPMENTAL STAGE: Expressed at higher levels from 12.5 dpc to
 15.5 dpc with a peak at 13.5 dpc. Expression in the developing
 liver as well as a restricted set of embryonic epithelial cells of
 the nasal cavity and pharyngo-tympanic tubes.
 CC -!- SIMILARITY: Belongs to peptidase family S1.
 CC -!- SIMILARITY: Contains 2 CUB domains.
 CC -!- SIMILARITY: Contains 3 LDL-receptor class A domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 or send an email to license@sib-sib.ch.

 CC EMBL; AY240929; AAP69827.1; -;
 CC EMBL; AK004939; BAB23684.2; -;
 CC EMBL; BC029645; AAB29645.2; -;
 CC HSRF; P00763; IDPO.
 CC MEROPS; S01.308; -;
 CC MGD; MGI:1919003; TMPRSS6.
 CC InterPro; IPR000859; CUB.
 CC InterPro; IPR009003; Cys Ser trypsin.
 CC InterPro; IPR002172; LDL_receptor_A.
 CC InterPro; IPR001254; Peptidase_S1.
 CC InterPro; IPR001314; Peptidase_S1a.
 CC Pfam; PF00057; ldl_recept_a; 2.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00361; LDLRECEPTOR.
 CC PROSITE; PS01180; CUB; 1.
 CC PROSITE; PS01209; LDLRA_1; 1.
 CC PROSITE; PS00068; LDLRA_2; 3.
 CC PROSITE; PS00240; TRYPsin_DOM; 1.
 CC PROSITE; PS00134; TRYPsin_HIS; 1.
 CC PROSITE; PS00135; TRYPsin_SER; 1.
 CC Hydrolase; Serine protease; Repeat; Signal-anchor; Transmembrane;
 CC Glycoprotein.
 CC DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 60 80 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 (POTENTIAL).
 CC DOMAIN 81 811 EXTRACELLULAR (POTENTIAL).
 CC DOMAIN 213 336 CUB 1.
 CC DOMAIN 323 440 CUB 2.
 CC DOMAIN 445 477 LDL-RECEPTOR CLASS A 1.
 CC DOMAIN 478 514 LDL-RECEPTOR CLASS A 2.
 CC DOMAIN 518 555 LDL-RECEPTOR CLASS A 3.
 CC DOMAIN 565 799 SERINE PROTEASE.
 CC ACT SITE 617 617 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC ACT SITE 668 668 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC ACT SITE 762 762 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC CARBOHYD 138 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 690 P -> PP (IN REF. 2).
 SQ SEQUENCE 811 AA; 90978 MW; 32EB3E7C3127801B CRC64;
 Query Match 18.4%; Score 415.5; DB 1; Length 811;
 Best Local Similarity 37.0%; Pred. No. 8.2e-26;
 Matches 102; Conservative 47; Mismatches 90; Indels 37; Gaps 11;
 QY 130 DCADG--KLKPCQCKQLRPRFKIIGFETIENQPFWFAAIYRHRGSGVTVYCGSLIS 187
 Db DCRDGSDECHDCGLQGLSSR--IVGGTVSEGEWPAQSLQIRGR-----HICGGALIA 608
 QY 198 PCWISATHCIDYFKED-----YIVYLGRSLNSNTGEMKFEVENLIHKDYSAD 240
 Db DRWVITAAHCF-----QEDSNASPKLMTVFLGKQRQNSRWPGEVSVKVSRLFLHPYHEED 663
 QY 241 TLAHNDIALLKIRSKGRCAQP---SRTIQICLPMSYNDPQFGTSCETITFGKENSSTD 297
 Db --SHDYDVALQL-----DHPVVSATVPVCLPARSHFFEPQCHWITGWAQREGG 714
 QY 298 VLYPEQLKMTVVKLISHRECQPHYGVSEVTTMLCAADPQWKTDCQDGGGLVC-SL 356
 Db PV-SNTLOKVDVLPQDLCSEAYRY--QVSPRLCAGYRKGDACQDGGGLVCREP 771
 QY 357 QGRMTLTGIVSWGRCALKDKPGVYTVRSHPFWIR 392
 Db SCRWFAGLVNGLGCGRPNFFGVYTVRVINNIQ 807
 RESULT 38
 PSS8 HUMAN STANDARD; PRT; 343 AA.
 AC Q16651; Q9UCA3;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Protastin precursor (EC 3.4.21.-).
 GN PRS8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Prostate;
 RX MEDLINE=95286644; PubMed=7768952;
 RA Yu J.X., Chao L., Chao J.;
 RT "Molecular cloning, tissue-specific expression, and cellular
 localization of human protastin mRNA.";
 RL J. Biol. Chem. 270:13483-13489(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Shat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield A.S.N., Krzywinski M.I., Skalska U., Smalius D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002). [3]

SEQUENCE OF 45-64.

TISSUE=semen;

MEDLINE=94308140; PubMed=8034638;

Yu J.X., Chao L., Chao J.;

"Prostasin is a novel human serine proteinase from seminal fluid. Purification, tissue distribution, and localization in prostate gland."; J. Biol. Chem. 269:18843-18848(1994).

-!- FUNCTION: Possesses a trypsin-like cleavage specificity.

-!- SUBUNIT: Heterodimer of two chains, light and heavy, held by a disulfide bond.

-!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. SECRETED AFTER CLEAVAGE OF INS C-TERMINUS.

-!- TISSUE SPECIFICITY: Found in prostate, liver, salivary gland, kidney, lung, pancreas, colon, bronchus and renal proximal tubular cells. In the prostate gland it may be synthesized in epithelial cells, secreted into the ducts, and excreted into the seminal fluid.

-!- SIMILARITY: Belongs to peptidase family S1.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; L41351; AAC41759.1; -

EMBL; U33446; AAB19071.1; -

EMBL; BC001462; AAB01462.1; -

PIR; A57014; A57014.

HSSP; P00763; 1DPO.

MEROPS; S01.159; -

GeneW; HGNC:9491; PRSS8.

MIM; 600823; -

GO; GO:0005615; C:extracellular space; TAS.

GO; GO:0005886; C:plasma membrane; TAS.

GO; GO:0008236; F:serine-type peptidase activity; TAS.

InterPro; IPR009003; Cys Ser trypsin.

InterPro; IPR001254; Peptidase_S1.

InterPro; IPR001314; Peptidase_S1A.

Pfam; PF00089; trypsin; 1.

PRINTS; PR00722; CHYMOTRYPSIN.

SMART; SM00020; TRYPSIN; 1.

PROSITE; PS02040; TRYPSIN_DOM; 1.

PROSITE; PS00134; TRYPSIN_HIS; 1.

PROSITE; PS00135; TRYPSIN_SER; 1.

Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein; Transmembrane.

KEYWORD	1	29
SIGNAL	1	POTENTIAL.
PROPEP	30	ACTIVATION PEPTIDE.
CHAIN	33	PROSTATIN LIGHT CHAIN.
CHAIN	45	PROSTATIN HEAVY CHAIN.
PROPEP	323	
TRANSMEM	323	POTENTIAL.
DOMAIN	320	SERINE PROTEASE.
DISULFID	45	INTERCHAIN (BY SIMILARITY).
DISULFID	37	BY SIMILARITY.
DISULFID	70	BY SIMILARITY.
DISULFID	168	BY SIMILARITY.
DISULFID	201	BY SIMILARITY.
DISULFID	223	BY SIMILARITY.
DISULFID	234	BY SIMILARITY.
ACT_SITE	85	CHARGE RELAY SYSTEM.
ACT_SITE	134	CHARGE RELAY SYSTEM.
ACT_SITE	238	CHARGE RELAY SYSTEM.
CARBOHYD	159	N-LINKED (GLCNAC...) (POTENTIAL).

```

SQ SEQUENCE 343 AA; 36431 MW; 98D6447F5A8C1B2 CRC64;
Query Match 18.1%; Score 408.5; DB 1; Length 343;
Best Local Similarity 36.6%; Pred. No. 1.1e-25;
Matches 100; Conservative 36; Mismatches 100; Indels 37; Gaps 8;

QY 140 CQKTLRPRFKIIGGEFTTINQWFAAIVRRHGGSVTY----VCGGSLISPCWVTSAT 195
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 37 CG---VAPQARITGSSAVAGQWPQV-----SITVEGHVCGGSLVSEQWLSAA 84
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 196 HCFIDYPKKEDYIVYLGSRSLNSNTQGBMKFEVENLILHKDYSADTLAHHNDTALLKIRS 255
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 85 HCFPEHHKEAYEVKLGHAHQDSSEDAKVTSLKDIIPHSYLOE--GSGDIALQL-- 140
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 256 KEGRCAQPSRTIQTICLSMYNDPQFGTSCBITGFK-ENSTDYLYPEQLKMTVVKLISH 314
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 --SRPITTSRYIRPCLIPAAVASPENGHLHCTVTGHWVAPSVSLTTPKPLQQLLEVPLSR 198
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 315 REC-----QOPHYVGSEVTTKMLCAADPOKWTDCGDSGGGLVCSLQGRMTLTG 364
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 199 ETCNCLYNIDAKPEPFP-----VQEDMVVAGYEGGKDACQDSGGSPFCFVGLMYLTG 254
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 365 IVSWGRCALKDKPGVYTRVSHFLPWRISHTKE 397
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 255 IVSWGDCAGARNRPGVYTLASSYASWQSKVTE 287
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 39
KAL MOUSE STANDARD; PRT; 638 AA.
ID KAL MOUSE
AC P26262;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasma kallikrein precursor (EC 3.2.1.34) (Plasma prekallikrein)
DE (Kininogenin) (Fletcher factor).
GN KLKB1 OR KLB3 OR PK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=91090844; PubMed=2264928;
RA Seidah N.G., Sawyer N., Hamelin J., Mion P., Beaubien G.,
RA Brachpapa L., Rochement J., Maikay M., Chretien M.;
RT "Mouse plasma kallikrein: cDNA structure, enzyme characterization,
RT and comparison of protein and mRNA levels among species.";
RL DNA Cell Biol. 9:737-748(1990).
CC -!- FUNCTION: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It
CC activates, in a reciprocal reaction, factor XII after its binding
CC to a negatively charged surface. It also releases bradykinin from
CC HMW kininogen and may also play a role in the renin-angiotensin
CC system by converting prorenin into renin.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Xaa and Lys-|-Xaa
CC bonds, including Lys-|-Arg and Arg-|-Ser bonds in (human)
CC kininogen to release bradykinin.
CC -!- SUBUNIT: The zymogen is activated by factor XIIa, which cleaves
CC the molecule into a light chain, which contains the active site,
CC and a heavy chain, which associates with HMW kininogen. These
CC chains are linked by one or more disulfide bonds.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasma kallikrein
CC subfamily.
CC -!- SIMILARITY: Contains 4 apple domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage, by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch)

```

CC EMBL; M59588; AAA63393.1; --
DR PIR; A36557; KOMSPL.
DR HSP; P00750; IATF.
DR MEROPS; S01.212; --
DR MGD; MGI:102849; Kikbl.
DR InterPro; IPR00177; Apple.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00024; PAN; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00005; APPLIEDOMAIN.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SMO0223; APPLE; 4.
DR SMART; SMO0020; TRYP_SPC; 1.
DR PROSITE; PS00495; APPLE; 4.
DR PROSITE; PS0240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease; Glycoprotein; Plasma; Zymogen; Signal;
KW Fibrinolysis; Blood coagulation; Inflammatory response; Liver;
KW Repeat.
FT SIGNAL 1 19
FT CHAIN 20 390 PLASMA KALLIKREIN HEAVY CHAIN.
FT CHAIN 391 638 PLASMA KALLIKREIN LIGHT CHAIN.
FT DOMAIN 120 105 APPLE 1.
FT DOMAIN 110 195 APPLE 2.
FT DOMAIN 200 285 APPLE 3.
FT DOMAIN 291 376 APPLE 4.
FT DOMAIN 389 621 SERINE PROTEASE.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 453 453 O-LINKED (PROBABLE).
FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (PROBABLE).
FT ACT_SITE 434 434 CHARGE RELAY SYSTEM.
FT ACT_SITE 483 483 CHARGE RELAY SYSTEM.
FT ACT_SITE 578 578 CHARGE RELAY SYSTEM.
FT DISULFID 21 104 BY SIMILARITY.
FT DISULFID 47 77 BY SIMILARITY.
FT DISULFID 51 57 BY SIMILARITY.
FT DISULFID 111 194 BY SIMILARITY.
FT DISULFID 137 166 BY SIMILARITY.
FT DISULFID 141 147 BY SIMILARITY.
FT DISULFID 201 284 BY SIMILARITY.
FT DISULFID 227 256 BY SIMILARITY.
FT DISULFID 231 237 BY SIMILARITY.
FT DISULFID 292 375 BY SIMILARITY.
FT DISULFID 318 347 BY SIMILARITY.
FT DISULFID 322 328 BY SIMILARITY.
FT DISULFID 340 345 BY SIMILARITY.
FT DISULFID 383 503 BY SIMILARITY.
FT DISULFID 419 435 BY SIMILARITY.
FT DISULFID 517 584 BY SIMILARITY.
FT DISULFID 548 563 BY SIMILARITY.
FT DISULFID 574 602 BY SIMILARITY.
SQ SEQUENCE 638 AA; CC27C93AB1086599 CRC64;

ery Match 18.1%; Score 407.5; DB 1; Length 638;
st Local Similarity 33.1%; Pred No. 2.7e-25;
tches 93; Conservative 56; Mismatches 107; Indels 25; Gaps 8;

QY 126 CMVHDCAGKLFQCGQKTLPRKTIIGGETTIENOPWFAIYRRHGGSVTVCGSL 185
Df 375 CKLVDSPD-----C---TTKINARIVGTTNAGLGPWQVSL--QVKLVSTQLCGGSI 423
QY 186 ISPCWVISAHCDFIDPKKEDYIVYLGSRSLNSQTGEMKPEVENLILHKDYSDATLAHH 245
Df 424 IGRQWLPAACFGDIPFDVWRVYGGILSLSEITKETPSSRIKELIHQEKYS--EGN 481
QY 246 NDIALKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFG---KENSTDYLYPE 302

Db 482 YDIALIKLQTP-----LNYTFQKPICLPSKADTNTIYTNCWVITGWGTYKEQGETQNI--- 534
QY 303 QLKMTVVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDSDGGLVCSLQGRWTL 362
Db 535 -LQKATPLVNEECQK-KYRDYVINKQMCAGYKEGTDACKDSDGGLVCKHSGRWQL 592
QY 363 TGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHYKKEENGLAL 403
Db 593 VGIITSWGEGCGRKDQPGVYTKVSEYMDWILEKTQSSDVRL 633

RESULT 40
TMSS_MOUSE
ID TMSS_MOUSE STANDARD; PRT; 455 AA.
AC Q9ER04; Q9ER02; Q9ER03;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transmembrane protease, serine 5 (EC 3.4.21.-) (Spinesin).
GN TMPS5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE=Brain;
RA Mitsui S., Yamaguchi N.;
RT "cDNA cloning of mouse spinesin."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 4).
RC TISSUE=Brain;
RA Mitsui S., Yamaguchi N.;
RT "Molecular cloning of mouse type 4 spinesin."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -! SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -! ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=4;
CC Name=4;
CC IsoId=Q9ER04-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9ER04-2; Sequence=VSP_005397, VSP_005398;
CC Name=2;
CC IsoId=Q9ER04-3; Sequence=VSP_005395;
CC Name=3;
CC IsoId=Q9ER04-4; Sequence=VSP_005396;
CC -! SIMILARITY: Belongs to peptidase family S1.
CC -! SIMILARITY: Contains 1 SRCR domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AB016229; BAB20276.1; --
DR EMBL; AB016230; BAB20277.1; --
DR EMBL; AB016423; BAB20278.1; --
DR EMBL; AB041037; BAB40328.1; --
DR HSSP; P00763; IDPO.
DR MEROPS; S01.313; --
DR MGD; MGI:1933407; Tmpres5.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SMO0020; TRYP_SPC; 1.

ULT 41
D HUMAN

FT ACT_SITE 218 218 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 42 178 BY SIMILARITY.
 FT DISULFID 61 177 BY SIMILARITY.
 FT DISULFID 157 224 BY SIMILARITY.
 FT DISULFID 189 203 BY SIMILARITY.
 FT DISULFID 214 239 BY SIMILARITY.
 FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 170 180 VNYFKTLQCAN -> GHHPHRWPEAP (IN REF. 3).
 SQ SEQUENCE 277 AA; 30570 MW; BAB9B8DFB5D542 CRC64;
 seq Match 17.9%; Score 405; DB 1; Length 277;
 st Local Similarity 41.1%; Pred. No. 1.7e-25;
 ches 101; Conservative 34; Mismatches 81; Indels 30; Gaps 10;
 Qy 153 GGEFTTTEPFAAIYRRHGGSVTVCGGSLISPCWVISAETHCIDIPKEDYIVYLG 212
 Db 38 GGYTCFPHSQPWAALVQGR-----LLCGVLVHPKWLTAACHL-----REGKLVYLG 87
 Qy 213 RSLRNSNTQCEMKFEVENLILKDY--SADTLAHHNDIALKRSKRGCAQPSRTIQT 270
 Db 88 KHALGRVEAGEQVREVHSPHPEYRSPTHLNHDHIMLELQSP-----VLTGYIQT- 142
 Qy 271 CLPSMYND---PQGTSCBITGFKENSTDYLPQKMTVVKLISHRECCQPHYGVSEV 327
 Dt 143 -LPLSHNNRLTP--GTTCTRVSGWGTTSPOVNPYKTLQCANIQLSDECRQ--VYFGKI 197
 Qy 328 TTKMLCAADPOWTKDSQGGSGGLVCSLQGRMTLGTIVSGWR-GCALKDXPGVYTRVSH 386
 Db 198 TDNMLCAGTRGGKDCGGSGGLVGN-----RTLYGIVSGWDFPGCQPPRGVYTRVSR 253
 Qy 387 FLPMIR 392
 Dt 254 YVLWIR 259
 RE JT 42
 HE HUMAN
 ID HEPES HUMAN STANDARD; PRT; 417 AA.
 AC P05981;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serine protease hepsin (EC 3.4.21.-) (Transmembrane protease, serine
 1).
 GN HPN OR TMPSR81.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RI SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=88209431; PubMed=2835076;
 RT Leytus S.P., Loeb K.R., Hagen F.S., Kurachi K., Davie E.W.;
 R1 "A novel trypsin-like serine protease (hepsin) with a putative
 transmembrane domain expressed by human liver and hepatoma cells.";
 RL Biochemistry 27:1067-1074(1988).
 RN [2]
 RE SEQUENCE FROM N.A.
 RC TISSUE=Pancreas, and Spleen;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalley D.E.,
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=91358502; PubMed=1885621;
 RA Tsuji A., Torres-Rosado A., Arai T., le Beau M.M., Lemons R.S.,
 Chou S.H., Kurachi K.;
 "Hepsin, a cell membrane-associated protease. Characterization,
 tissue distribution, and gene localization.";
 J. Biol. Chem. 266:16948-16953(1991).
 RL [4]
 RP CHARACTERIZATION.
 RX MEDLINE=93348237; PubMed=8346233;
 RA Torres-Rosado A., O'Shea K.S., Tsuji A., Chou S.H., Kurachi K.;
 "Hepsin, a putative cell-surface serine protease, is required for
 mammalian cell growth.";
 Proc. Natl. Acad. Sci. U.S.A. 90:7181-7187(1993).
 RL -!- FUNCTION: Plays an essential role in cell growth and maintenance
 of cell morphology.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
 CC -!- TISSUE SPECIFICITY: Present in most tissues, with the highest
 level in liver.
 CC -!- SIMILARITY: Belongs to peptidase family S1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.iesb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL; M18930; AAA36013.1; .
 DR EMBL; X07732; CAA30558.1; .
 DR EMBL; X07002; CAA30058.1; .
 DR EMBL; BC025716; AAB25716.1; .
 DR F01; S00845; S00845.
 DR HSP; P00763; IDPO.
 DR MEROPS; S01.224; .
 DR Genew; HGNC:5155; HPN.
 DR MIM; 142440; .
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0008236; F:serine-type peptidase activity; TAS.
 DR GO; GO:0008151; P:cell growth and/or maintenance; TAS.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYPSIN_DOM; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Transmembrane; Signal-anchor.
 FT CHAIN 1 162
 FT SERINE PROTEASE HEPSPIN, NON-CATALYTIC
 FT CHAIN (POTENTIAL).
 FT SERINE PROTEASE HEPSPIN, CATALYTIC CHAIN
 FT (POTENTIAL).
 FT CYTOPLASMIC (POTENTIAL).
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT (POTENTIAL).
 FT EXTRACELLULAR (POTENTIAL).
 FT SERINE PROTEASE.
 FT CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 257 257
 FT ACT_SITE 353 353

```
F DISULFID 153 277 INTERCHAIN (BY SIMILARITY).
F DISULFID 188 204 BY SIMILARITY.
F DISULFID 322 338 BY SIMILARITY.
F DISULFID 349 349 BY SIMILARITY.
F CARBOHYD 112 112 N-LINKED (GLCNAC...) (POTENTIAL).
F SEQUENCE 417 AA; B2086FF661B551D7 CRC64;
      Query Match 17.9%; Score 404.5; DB 1; Length 417;
      1st Local Similarity 34.9%; Pred. No. 2.9e-25;
     atches 106; Conservative 55; Mismatches 98; Indels 45; Gaps 15;
C 122 LVQECMVHDCADGK-----LRFQCQKQLRPRFKIIGGFTTINQWFAIYRRHGGSV 177
I 131 LLEVISVCDPCRGFLAAICQDCGRKL-PVDRIVGGGRDTSIGRWPMQVSL--RYDG--- 184
C 178 TYVCGSLSPCWISATHCFIDYPKEDYIVLGRSLNSNTQGE-----MKFEVENLI 232
I 185 AHLGCGSLSGDWLTAAHCF-----PERNR---VLSRWRFVAGAAQSPHGLQGVQAAV 238
C 233 LHKDY-----SADTLAHNDIALKIRSGKRCQAPSRITQICLPMSYNDPQGTGCEIT 288
I 239 YHGGLYFPDRPNSEENDIALVHLSF-----LPLTEYIQVCLPAAGQALVDGKICTVT 294
C 289 GFGENSTDYLPDQ-----LQTVVKLISHRECQPHYGVSEVTTKMLCAADPQWKDSCQ 345
I 295 GWG---NTQY-YGQAGVLOEARPIISNDVCGADFYGNQIKRPMFCAGYPEGGDACQ 350
C 346 GDSGGPLVVC---SLQGRMTLTGVSWGRGALKDKPVTVTRVSHLPW---IRGHTKE 397
I 351 GDSGGPFVCEDSISRTPRWRLGVISWGTGTCALAKQKGVTVTKVSDPREMIFQAIKTHS-E 409
C 398 ENGL 401
I 410 ASGM 413
      ULT 43
      _HUMAN
      KAL_HUMAN STANDARD; PRT; 638 AA.
      P03952;
      23-OCT-1986 (Rel. 02, Created)
      23-OCT-1986 (Rel. 02, Last sequence update)
      15-MAR-2004 (Rel. 43, Last annotation update)
      Plasma kallikrein precursor (EC 3.4.21.34) (Plasma prekallikrein)
      (kininogenin) (Fletcher factor).
      KLKB1 OR KLK3.
      Homo sapiens (Human).
      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
      NCBI_TaxID=9606;
      [1]
      SEQUENCE FROM N.A.
      MEDLINE=86243359; PubMed=3521732;
      Chung D.W., Fujikawa K., McMullen B.A., Davie E.W.;
      "Human plasma prekallikrein, a zymogen to a serine protease that
      contains four tandem repeats.";
      Biochemistry 25:2410-2417(1986).
      [2]
      SEQUENCE FROM N.A. AND VARIANTS SER-143; GLN-202 AND PRO-208.
      MEDLINE=20487549; PubMed=11031105;
      Yu H., Anderson P.J., Freedman B.I., Rich S.S., Bowden D.W.;
      "Genomic structure of the human plasma prekallikrein gene,
      identification of allelic variants, and analysis in end-stage renal
      disease.";
      Genomics 69:225-234(2000).
      [3]
      SEQUENCE FROM N.A., AND VARIANTS SER-143; THR-178; GLN-202; CYS-269;
      VAL-311; ALA-358; ALA-381; PRO-442 AND GLN-560.
      Rieder M.J., Armet T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
      Rajkumar N., Toch E.J., Yi Q., Nickerson D.A.;
      Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
      [4]
      PARTIAL SEQUENCE, AND DISULFIDE BONDS.
```

```
RX MEDLINE=91152016; PubMed=1998666;
RA McMullen B.A., Fujikawa K., Davie E.W.;
RT "Location of the disulfide bonds in human plasma prekallikrein: the
RT presence of four novel apple domains in the amino-terminal portion of
RT the molecule.";
RL Biochemistry 30:2050-2056(1991).
RN [5]
RP CARBOHYDRATE-LINKAGE SITE ASN-453.
RX MEDLINE=22660472; PubMed=12754519;
RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
RT "Identification and quantification of N-linked glycoproteins using
RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";
RL Nat. Biotechnol. 21:660-666(2003).
CC -!- FUNCTION: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It
CC activates, in a reciprocal reaction, factor XII after its binding
CC to a negatively charged surface. It also releases bradykinin from
CC HMW kininogen and may also play a role in the renin-angiotensin
CC system by converting prorenin into renin.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Xaa and Lys-|-Xaa
CC bonds, including Lys-|-Arg and Arg-|-Ser bonds in (human)
CC kininogen to release bradykinin.
CC -!- SUBUNIT: The zymogen is activated by factor XIIa, which cleaves
CC the molecule into a light chain, which contains the active site,
CC and a heavy chain, which associates with HMW kininogen. These
CC chains are linked by one or more disulfide bonds.
CC -!- DISEASE: Defects in KLKB1 are the cause of Fletcher factor
CC deficiency (MIM:229000); a blood coagulation defect.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasma kallikrein
CC subfamily.
CC -!- SIMILARITY: Contains 4 apple domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M13143; AAA60153.1; -.
CC EMBL; AF232742; AAF79940.1; -.
CC EMBL; AF232734; AAF79940.1; JOINED.
CC EMBL; AF232735; AAF79940.1; JOINED.
CC EMBL; AF232736; AAF79940.1; JOINED.
CC EMBL; AF232737; AAF79940.1; JOINED.
CC EMBL; AF232738; AAF79940.1; JOINED.
CC EMBL; AF232739; AAF79940.1; JOINED.
CC EMBL; AF232740; AAF79940.1; JOINED.
CC EMBL; AF232741; AAF79940.1; JOINED.
CC EMBL; AF232741; AAF79940.1; JOINED.
CC EMBL; AV190920; AAN84794.1; -.
CC PIR; A00921; KQHUP.
CC HSP; P00763; IDPO.
CC MEROPS; S01.212; -.
CC Genew; HGNC:6371; KLKB1.
CC MIM; 229000; -.
CC GO; GO:0003807; F:plasma kallikrein activity; TAS.
CC GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
CC InterPro; IPR000177; Apple.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR003014; PAN.
CC InterPro; IPR001254; Peptidase S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00024; PAN; 4.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PRO0005; APLEDOMAIN.
CC PRINTS; PRO0722; CHYMOTRYPSIN.
CC SMART; SM00223; APPLE; 4.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS00495; APPLE; 4.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Hydrolase; Serine protease; Glycoprotein; Plasma; Zymogen; Signal;
KW
```


"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [6]
 Cao J., Fan W., Zheng S.;
 "Genomic analysis of a novel human serine protease SNCI19.";
 Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 [7]
 CHARACTERIZATION.
 TISSUE=Milk;
 MEDLINE=99303582; PubMed=10373425;
 Lin C.Y., Anders J., Johnson M., Dickson R.B.;
 "Purification and characterization of a complex containing matrixinase
 and a Kunitz-type serine protease inhibitor from human milk.";
 J. Biol. Chem. 274:18237-18242(1999).
 C -1- FUNCTION: Degrades extracellular matrix. Proposed to play a role
 in breast cancer invasion and metastasis. Exhibits trypsin-like
 activity as defined by cleavage of synthetic substrates with Arg
 or Lys as the P1 site.
 C -1- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
 C -1- SIMILARITY: Belongs to peptidase family S1.
 C -1- SIMILARITY: Contains 2 CUB domains.
 C -1- SIMILARITY: Contains 4 LDL-receptor class A domains.
 C This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 D EMBL; AF118224; AD42765.2; -
 D EMBL; AF133086; AAF00109.1; -
 D EMBL; AB030036; BAB20376.1; -
 D EMBL; AF057145; AAG15395.1; -
 D EMBL; BC005826; AAH05826.1; -
 D EMBL; BC030532; AAH30532.1; -
 D EMBL; AF283256; AAG13949.1; -
 D HSP; P00763; IDPO.
 D Genew; HGNC:11344; ST14.
 D MIM; 606797; -
 D MEROPS; S01.302; -
 D GO; GO:0005887; C: integral to plasma membrane; TAS.
 D GO; GO:0008236; F: serine-type peptidase activity; TAS.
 D GO; GO:0006508; P: proteolysis and peptidolysis; TAS.
 D InterPro; IPR000859; CUB.
 D InterPro; IPR009003; Cys_Ser_trypsin.
 D InterPro; IPR02172; LDL_receptor_A.
 D InterPro; IPR01254; Peptidase_S1.
 D InterPro; IPR01314; Peptidase_S1A.
 D Pfam; PF00431; CUB_2
 D Pfam; PF00057; ldl_recept_a; 4.
 D Pfam; PF00089; trypsin; 1.
 D PRINTS; PR00722; CHYMOTRYPSIN.
 D PRINTS; PR00261; LDLRECEPTOR.
 D SMART; SM00042; CUB; 2.
 D SMART; SM00192; LDLA; 3.
 D SMART; SM00020; TRYD_SPC; 1.
 D PROSITE; PS01180; CUB_2.
 D PROSITE; PS01209; LDLRA_1; 2.
 D PROSITE; PS00068; LDLRA_2; 4.
 D PROSITE; PS00240; TRYPSIN_DOM; 1.
 D PROSITE; PS00134; TRYPSIN_HIS; 1.
 D PROSITE; PS00135; TRYPSIN_SER; 1.
 D Signal-anchor; Glycoprotein; Hydrolase; Serine protease;
 D Transmembrane; Repeat.
 D DOMAIN 1 55 CYTOPLASMIC (POTENTIAL).
 D TRANSMEM 56 76 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 D DOMAIN 77 855 (POTENTIAL).
 D DOMAIN 214 334 EXTRACELLULAR (POTENTIAL).
 D CUB 1.

FT DOVAIN 340 447 CUB 2.
 FT DOVAIN 452 487 LDL-RECEPTOR CLASS A 1.
 FT DOVAIN 487 524 LDL-RECEPTOR CLASS A 2.
 FT DOVAIN 524 560 LDL-RECEPTOR CLASS A 3.
 FT DOVAIN 566 603 LDL-RECEPTOR CLASS A 4.
 FT DOVAIN 615 854 SERINE PROTEASE.
 FT ACT SITE 656 656 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 711 711 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 805 805 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 485 485 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 772 772 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 327 329 PEA -> GTR (IN REF. 5; AAH05826).
 FT CONFLICT 381 381 R -> S (IN REF. 4).
 FT CONFLICT 674 674 A -> V (IN REF. 3).
 SQ SEQUENCE 855 AA; 94769 MW; 26143132C01F99C9 CRC64;
 Query Match 17.8%; Score 401.5; DB 1; Length 855;
 Best Local Similarity 28.1%; Pred. No. 1.2e-24;
 Matches 119; Conservative 65; Mismatches 157; Indels 83; Gaps 20;
 QY 5 HOVPSNCDCLNGG--TCVSNKYPSNIHW-----CNCPK---KEGGOHCE 43
 DB 480 HSDLNLSCDAGHQFTC-KKKCKPLFWVCDSDVNDGDSDEQCSCPAQTFRCSNGKC- 537
 QY 44 IDKSKTCYEGNGHFYRGKASTDTMGRCPLPNSATVLQOTYHAHRSALQGLGKHYCR 103
 DB 538 LSKSQQC---NGKDDCGGSDGA--SCPKVNVVCTKHTYRCLNGLCLSGK----- 583
 QY 104 NPDNRPRWCYVQGLKPLVQEC--WVHDCADG--KLKFCQCGKTLRPRFKLIGGEFTTIE 160
 DB 594 NP-----ECGKEDCDSDGDEKDCDCGLRSFTQARVVGVTGDADEG 624
 QY 161 NQPFAAIYRHRGSGVTVYCGGSLISPCWVISATHCFID-----YPKKEDIYVLG-RS 214
 DB 625 EWPQVSLHALGQG---HICGASLISPNMLVSAACHVIDDRGRFYSDPTQWTAFLGHD 680
 QY 215 RLNSNTQGMKFEVENILHKDYADTLAHNDTALKIRSKGRCQAQPSFTICILPS 274
 DB 681 QSRSAPGVQERRLKRITISHPFNDFTDY--DIALLELE---KPAEYSMWVRPILCPD 734
 QY 275 MYNDPQFGTSCIEITGFG--KENSTDYLYPEQLKVTWVVKLISHRECQCPHYVGSVTKML 332
 DB 735 ASHVFPAGKAIWTVGWGHTVQGTGALI---LQKGEIRVINQTTCE--NLLPQGITPRMM 789
 QY 333 CAADPQWKTDSCQDSDGGL--VCSLQGRMLTGVSWGRGALKDKPKGVTVRVSHFLPWI 391
 DB 790 CVGFLSGGVSDSCQDSDGGL--SSVEADGRIFQAGVYVSWGDCGAQRNKPQVYTRLPFLFRDWI 849
 QY 392 RSHT 395
 DB 850 KENT 853
 RESULT 45
 ST14 MOUSE
 ID ST14 MOUSE STANDARD; PRN; 855 AA.
 AC P56677;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Suppressor of tumorigenicity 14 (EC 3.4.21.-) (Epithin).
 GN ST14 OR PRSS14.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C.B.17SCID; TISSUE=Thymus;
 RX MEDLINE=99216440; PubMed=10199918;
 RA Kim M.G., Chen C., Lyu M.S., Cho E.G., Park D., Kozak C.,

RA Schwartz R.H.;
 RT "Cloning and chromosomal mapping of a gene isolated from thymic
 RT stromal cells encoding a new mouse type II membrane serine protease,
 RT epithin, containing four LDL receptor modules and two CUB domains."
 RL Immunogenetics 49:420-428 (1999).
 RN [2]
 RP REVISONS: TO 23; 321; 325; 343; 409-410 AND C-TERMINUS.
 RC STRAIN=C.B.17SCID; TISSUE=Thymus;
 RA Kim M.G., Chen C., Cho E.G., Park D., Schwartz R.H.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast; tumor;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
 RA Hopkins L.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RI human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RC -1- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
 CC -1- TISSUE SPECIFICITY: Highly expressed in intestine, kidney, lung,
 CC and thymus. Not expressed in skeletal muscle, liver, heart,
 CC testis and brain.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 2 CUB domains.
 CC -1- SIMILARITY: Contains 4 LDL-receptor class A domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DF EMBL; AF042822; AAD02230.3; -;
 DR EMBL; BC005496; AA05496.1; -;
 DR HSSP; P20231; 1AAO.
 DR MEROPS; S01.302; -;
 DR MGD; MGI:1338881; Sc14.
 DR GO; GO:0005576; C:extracellular; IDA.
 DR GO; GO:0019897; C:extrinsic to plasma membrane; IDA.
 DF GO; GO:0008236; F:serine-type peptidase activity; IDA.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR InterPro; IPR001254; Peptidase_S1.
 DF InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00431; CUB; 2.
 DR Pfam; PF00057; ldl_recept a; 4.
 DF Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR SMART; SM00042; CUB; 2.
 DF SMART; SM00192; LDLa; 4.
 DF SMART; SM00020; Tryp Spc; 1.
 DF PROSITE; PS01180; CUB; 2.
 DF PROSITE; PS01209; LDLRA_1; 2.

DR PROSITE; PS00068; LDLRA_2; 4.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Signal-anchor; Glycoprotein; Hydrolase; Serine protease;
 KW Transmembrane; Repeat.
 FT DOMAIN 1 55
 FT TRANSMEM 56 76
 FT
 FT DOMAIN 77 855
 FT DOMAIN 214 331
 FT DOMAIN 340 444
 FT DOMAIN 451 488
 FT DOMAIN 489 522
 FT DOMAIN 523 561
 FT DOMAIN 565 604
 FT DOMAIN 615 854
 FT ACT_SITE 656 656
 FT ACT_SITE 711 711
 FT ACT_SITE 805 805
 FT CARBOHYD 107 107
 FT CARBOHYD 302 302
 FT CARBOHYD 365 365
 FT CARBOHYD 421 421
 FT CARBOHYD 489 489
 FT CARBOHYD 772 772
 SQ SEQUENCE 855 AA; 94654 MW; 4F10B84DA2146DD5 CRC64;
 Query Match 17.8%; Score 401.5; DB 1; Length 855;
 Best Local Similarity 27.9%; Pred. No. 1.2e-24;
 Matches 117; Conservative 68; Mismatches 147; Indels 87; Gaps 20;
 QY 11 CDCLNGGTCVSNKYPSNIHW-----CNCPKFGQHC-----IDKSKTC 50
 DB 488 CNATHQFTC-KNQCFKPLFWCDSVNDGSDGSDGSCCPA--GSFKCSNGKCLPQSKC 544
 QY 51 YEGNGHFYRGKASITDTMGRPCLPWNSATVLOQTQYHAHRSDALQGLGKHNYCRNPDNR 110
 DB 545 ---NGKDCGSDGDEA---SCDSVNVVSTKYVRCQNGCLSKG-----NP----- 585
 QY 111 FWCYVQVGLKPLVQEC-MVHDCADG--KLKFCQCGKTLRPFKLIIGGFTTIENQPF 167
 DB 586 -----ECGKTCDSGSDGKNCDCGLRSTKQARVVGGTNADEGEFQVWS 631
 QY 168 IYRRHRGGSVTVCVCGSLSPCWVISATHCFIDVP--KKEDYIVY-----LGRSLNSN 219
 DB 632 LHALQGG---HLCCASLISPDWLVAACQDDKNFKYSDYTMWTAFLGLLDQSKRSAS 687
 QY 220 TQGMKFEVENILHKOYSADTLAHNDIALKIRSKGRCQAQPSRTTQICLPSNYNDP 279
 DB 688 --GVQELKLRITHPSPNDFDY--DIALLELE---KSVEXTVVRPICLPDATHVF 739
 QY 280 QFGTSCETIGFG--KENSTDYLYPEQLKMTVVKLISHRECQPHYVYGSEVTKMLCAADP 337
 DB 740 PAGKAIWTVGWGHTKEGGTGALI---LQGEIRVINQTTCD--LMPQOITPRMVCVGL 794
 QY 338 QWKTSQCGSDSGPL-VCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHLFWIRSH 395
 DB 795 SGGVSDSCQSDGSPLSAEDKGRMPQAGVSWGEGCAQRNKPFGVYTRLVVVRDVIKEHT 853
 RESULT 46
 HEPS_RAT STANDARD; PRT; 416 AA.
 ID HEPS_RAT
 AC Q05511;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serine protease hepsin (EC 3.4.21.-).
 GN HEPN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
TISSUE=Liver;
MEDLINE=93305733; PubMed=8318546;
Parley D., Reymond P., Nick H.;
"Cloning and sequence analysis of rat hepsin, a cell surface serine
protease";
Biochim. Biophys. Acta 1173:350-352(1993).
-!- FUNCTION: Plays an essential role in cell growth and maintenance
of cell morphology.
-!- SUBCELLULAR LOCATION: Type II membrane family S1.
-!- SIMILARITY: Belongs to peptidase family S1.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X70900; CAA50256.1; -;
PIR; S33777; S33777.
HSSP; P00763; IDPO.
MEROPS; S01.224; -;
InterPro; IPR009003; Cys Ser trypsin.
InterPro; IPR001254; Peptidase_S1.
InterPro; IPR001314; Peptidase_S1A.
InterPro; IPR001190; Srcr_receptor.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
SMART; SM00202; SR; 1.
SMART; SM00202; Tryp_Spc; 1.
PROSITE; PS00240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
Hydrolase; Serine protease; Transmembrane; Signal-anchor.
SERINE PROTEASE HEPsin, NON-CATALYTIC
CHAIN 1 161
CHAIN 162 416
CHAIN 162 416
DOMAIN 1 16
TRANSMEM 17 43
DOMAIN 44 416
DOMAIN 162 416
ACT_SITE 202 202
ACT_SITE 256 256
ACT_SITE 352 352
DISULFID 152 276
DISULFID 187 203
DISULFID 321 337
DISULFID 348 380
CARBOHYD 111 111
SEQUENCE 416 AA; 44926 MW; 85A9F8FA9550E180 CRC64;
very Match 17.8%; Score 401; DB 1; Length 416;
est Local Similarity 34.2%; Pred. No. 5.6e-25;
atches 109; Conservative 55; Mismatches 103; Indels 52; Gaps 17;

112 WCYQVGLKPLVQECM-----VHDCADGKLF-----CQSQKTLPRFKIIGGEFTTENOP 163
117 FVDEGGL-PLAQLLDVISVCDPCRFITATQCDCGRKL-PVDRIVGQSSSLGRWP 174
164 WFAAYRRHRGGSYYVCGSLISPCWVIGATHCFIDYPKKEYI-----VYLGRSLNSN 219
175 WQVSL--RYDG---THLCGSLSGDWLTAACHF-----PERNVLRSWRVFAG--AVART 224
220 TQGEKKEVENLILH-----KQYSADTLAHNDIALKIRSEGRCAQSRRTQICLP 273
225 SPHAVQLGVQAVIYHGGVLPFRDPTID--ENSNDIALVHLSSS-----LPLTEYIQVCLP 278

Qy 274 SMYNDPQFGTSCETIGTKENSTDIYPEO---LKMVTWVKLISHRECQOPHYVYGSEVTK 330
Db 279 AAGQALVDGKVKCTVIGNTQ-----FYQQAUVLQEARVPIISNEVCNSEDFFGNQIKPK 334
Qy 331 MLCAADPQWKTDSQCGSDGGPLYC-----SLQGRMTLTIGVSGRGKALCKDKPGVYTRVSH 386
Db 335 MFCAGYPEGGIDACQGDGSGGHFVCEDRISGTSRWRLCGIVSGTGCALARKPGVYTKVID 394
Qy 387 FLPW-----IRSHYKEENGL 401
Db 395 FREFIFOAIKTHS-EATGM 412

RESULT 47
TRY3_CHICK
ID TRY3_CHICK STANDARD; PRT; 248 AA.
AC Q90629;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Trypsin II-P29 precursor (EC 3.4.21.4).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=95251611; PubMed=7733885;
RT Wang K., Gan L., Lee I., Hood L.E.;
RT "Isolation and characterization of the chicken trypsinogen gene
family.";
RL Biochem. J. 307:471-479(1995).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -!- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: High levels are seen in the pancreas while
lower levels are found in the liver, spleen and thymus.
CC -!- SIMILARITY: Belongs to peptidase family S1.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U15157; AAA79914.1; -;
PIR; S55066; S55066.
HSSP; P00763; IDPO.
MEROPS; S01.151; -;
InterPro; IPR009003; Cys Ser trypsin.
InterPro; IPR001254; Peptidase_S1.
InterPro; IPR001314; Peptidase_S1A.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
SMART; SM00202; Tryp_Spc; 1.
PROSITE; PS00240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;
Calcium-binding; Signal; Multigene family.
KW SIGNAL 1 16
FT PROPEP 17 25
FT CHAIN 26 248
FT ACT_SITE 65 65
FT METAL 77 77
FT METAL 79 79
FT METAL 82 82
FT METAL 82 82

FT METAL 87 87 CALCIUM (BY SIMILARITY).
 FT ACT_SITE 109 109 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 202 202 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 32 162 BY SIMILARITY.
 FT DISULFID 50 66 BY SIMILARITY.
 FT DISULFID 134 235 BY SIMILARITY.
 FT DISULFID 141 208 BY SIMILARITY.
 FT DISULFID 173 187 BY SIMILARITY.
 FT DISULFID 198 222 BY SIMILARITY.
 FT SITE 196 196 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 SC SEQUENCE 248 AA; 26622 MW; 55E16B07622B588E CRC64;
 ry Match 17.5%; Score 395.5; DB 1; Length 248;
 st Local Similarity 38.7%; Pred. No. 8.5e-25;
 ches 94; Conservative 40; Mismatches 84; Indels 25; Gaps 7;
 QY 150 KIIGGFTTIENQWFAAIVRRHGGSVTVVCGSLISPCWISATHCFIDYPKEDYIV 209
 DC 25 KIVGVCYCPHSPYQVSL-----NSGYHFCGSLNSQWLSAHCY-----KSRIOV 73
 QY 210 YLGRSLNSNTQGMKFEVENLHKDYADTIAHNDIALKIRSKRGCAQPSFTIQ 269
 DC 74 RLGSYNDIQDESEVRSVSSVIRHPKYSSITL--NNDIMLIKLAS-----AVEYSADIQ 127
 QY 270 ICLPSMNDPOFGNSCITGFGKNSDLYLPEQLKMTVVKLISHRECQPHYVYGEVTT 329
 DC 128 IALPS--SCAKAGTECLISGNTLSNGVYPELLOCLNAPILSDQCEA--YPGDITS 183
 QY 330 KMLCAADPQWKTDSCQDGSGLYVCSLQGRMTLTGIVSWGRGKALDKPKGVYTRVSHFLP 389
 DC 184 NMICVGFLEGGKDCQDGSQGVVNCGE---LQGVSWIGGKALGKPGVYTKVNCYVD 239
 QY 390 WIR 392
 DC 240 WIQ 242
 LT 48
 TR BOVIN
 IF TRV1 BOVIN STANDARD; PRT; 243 AA.
 AC P00760;
 DT 21-JUL-1986 (Rel. 01, Created)
 D1 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Trypsinogen, cationic precursor (EC 3.4.21.4) (Beta-trypsin)
 DE (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RE SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RA Okajima T., Maniwa M., Nagao S., Fujikawa H., Goto S.;
 RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
 RE SEQUENCE OF 15-243, AND DISULFIDE BONDS.
 RX MEDLINE=67168848; PubMed=5967094;
 RA Mikes O., Holeysovsky V., Tomasek V., Sorm F.;
 RT "Covalent structure of bovine trypsinogen. The position of the
 remaining amides".
 RL Biochem. Biophys. Res. Commun. 24:346-352(1966).
 RN [3]
 RE REVISIONS.
 RX MEDLINE=72035053; PubMed=4399051;
 RA Hartley B.S.;
 RT "Homologies in serine proteinases".
 RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:77-87(1970).
 RN [4]
 RE REVISIONS.
 RX MEDLINE=75146445; PubMed=1092332;
 RA Titani K., Ericsson L.H., Neurath H., Walsh K.A.;

RT "Amino acid sequence of dogfish trypsin.";
 RL Biochemistry 14:1358-1366(1975).
 RN [5]
 RX X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF CALCIUM-BINDING SITE.
 RA MEDLINE=76072097; PubMed=512;
 RT Bode W., Schwager P.;
 RT "The refined crystal structure of bovine beta-trypsin at 1.8-A
 resolution. II. Crystallographic refinement, calcium binding site,
 benzamide binding site and active site at pH 7.0.";
 RL J. Mol. Biol. 98:693-717(1975).
 RN [6]
 RX X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RA MEDLINE=77112431; PubMed=556951;
 RT Kosiakoff A.A., Chambers J.L., Kay L.M., Stroud R.M.;
 RT "Structure of bovine trypsinogen at 1.9-A resolution.";
 RL Biochemistry 16:654-664(1977).
 RN [7]
 RX DISULFIDE BONDS.
 RA MEDLINE=66079271; PubMed=5892911;
 RT Kauffman D.L.;
 RT "The disulfide bridges of trypsin.";
 RL J. Mol. Biol. 12:929-932(1965).
 CC -- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -- COFACTOR: Binds 1 calcium ion per subunit.
 CC -- SUBCELLULAR LOCATION: Extracellular.
 CC -- TISSUE SPECIFICITY: Synthesized in the acinar cells of the
 pancreas.
 CC -- PTM: AUTOCATALYTIC CLEAVAGE AFTER LYS-20 LEADS TO BETA-TRYPSIN BY
 RELEASING A TERMINAL HEXAPEPTIDE. SUBSEQUENT CLEAVAGE AFTER
 LYS-145 LEADS TO ALPHA-TRYPSIN. FURTHER CLEAVAGE AFTER LYS-190
 YIELDS PSUDOTRYPSIN. A CLEAVAGE MAY ALSO OCCUR AFTER ARG-119.
 CC -- SIMILARITY: Belongs to peptidase family S1.
 CC -- DATABASE: NAME=Worthington enzyme manual;
 CC WWW="http://www.worthington-biochem.com/TRY/".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D38507; BAA07516.1; -
 DR PDB; 1A07; 25-FEB-98.
 DR PDB; 1A0J; 14-OCT-98.
 DR PDB; 1A28; 13-JAN-99.
 DR PDB; 1BJU; 13-JAN-99.
 DR PDB; 1BTV; 13-JAN-99.
 DR PDB; 1BTP; 29-JAN-96.
 DR PDB; 1BFW; 15-OCT-95.
 DR PDB; 1BTX; 15-OCT-95.
 DR PDB; 1BTY; 15-OCT-95.
 DR PDB; 1BTZ; 15-OCT-95.
 DR PDB; 1CIN; 30-JAN-02.
 DR PDB; 1C10; 26-SEP-01.
 DR PDB; 1C1P; 26-SEP-01.
 DR PDB; 1C1Q; 26-SEP-01.
 DR PDB; 1C1R; 26-SEP-01.
 DR PDB; 1C1S; 26-JUL-00.
 DR PDB; 1C1T; 26-SEP-01.
 DR PDB; 1C2D; 26-SEP-01.
 DR PDB; 1C3E; 26-SEP-01.
 DR PDB; 1C3F; 26-SEP-01.
 DR PDB; 1C2G; 26-SEP-01.
 DR PDB; 1C2H; 26-SEP-01.
 DR PDB; 1C2I; 26-SEP-01.
 DR PDB; 1C2J; 26-SEP-01.
 DR PDB; 1C2K; 26-SEP-01.
 DR PDB; 1C2L; 26-SEP-01.
 DR PDB; 1C2M; 26-SEP-01.
 DR PDB; 1C3P; 26-SEP-01.
 DR PDB; 1C5Q; 26-SEP-01.

D	PDB; 1C5R; 26-SEP-01.	DR	PDB; 1TLD; 15-OCT-92.
L	PDB; 1CSS; 26-SEP-01.	DR	PDB; 1TNG; 30-NOV-94.
D	PDB; 1C5T; 26-SEP-01.	DR	PDB; 1TNH; 30-NOV-94.
D	PDB; 1CSU; 26-SEP-01.	DR	PDB; 1TNI; 30-NOV-94.
D	PDB; 1CSV; 26-SEP-01.	DR	PDB; 1TNJ; 30-NOV-94.
L	PDB; 1CES; 23-MAR-99.	DR	PDB; 1TNK; 30-NOV-94.
D	PDB; 1CU8; 26-SEP-01.	DR	PDB; 1TNL; 30-NOV-94.
D	PDB; 1CU9; 26-SEP-01.	DR	PDB; 1TPA; 14-MAR-85.
D	PDB; 1D6R; 05-MAY-00.	DR	PDB; 1TPO; 14-MAR-85.
L	PDB; 1EB2; 11-MAR-02.	DR	PDB; 1TPP; 16-APR-87.
D	PDB; 1EBZ; 25-OCT-00.	DR	PDB; 1TPS; 26-JAN-95.
D	PDB; 1F0T; 20-SEP-00.	DR	PDB; 1TYN; 26-JAN-95.
D	PDB; 1F0U; 20-SEP-00.	DR	PDB; 1XUF; 16-DEC-98.
D	PDB; 1F2S; 05-JUN-00.	DR	PDB; 1XUG; 16-DEC-98.
D	PDB; 1G3S; 23-OCT-01.	DR	PDB; 1XUH; 11-NOV-98.
D	PDB; 1G3R; 17-JAN-01.	DR	PDB; 1XUI; 11-NOV-98.
D	PDB; 1G3D; 17-JAN-01.	DR	PDB; 1XUJ; 11-NOV-98.
D	PDB; 1G3E; 17-JAN-01.	DR	PDB; 1XUK; 11-NOV-98.
D	PDB; 1G9I; 11-APR-01.	DR	PDB; 1YYI; 08-JUN-99.
D	PDB; 1GBT; 31-JAN-94.	DR	PDB; 2BTC; 19-JAN-00.
D	PDB; 1GHZ; 22-JAN-02.		
D	PDB; 1GI0; 22-JAN-02.		
D	PDB; 1GI2; 22-FEB-02.		
D	PDB; 1GI3; 22-JAN-02.		
D	PDB; 1GI4; 22-FEB-02.		
D	PDB; 1GI5; 22-FEB-02.		
D	PDB; 1GI6; 22-JAN-02.		
D	PDB; 1GJ6; 27-APR-02.		
D	PDB; 1J8A; 12-SEP-01.		
D	PDB; 1JIR; 18-JUL-01.		
D	PDB; 1JRS; 14-OCT-96.		
D	PDB; 1JRT; 14-OCT-96.		
D	PDB; 1K1I; 28-NOV-01.		
D	PDB; 1K1J; 28-NOV-01.		
D	PDB; 1K1L; 28-NOV-01.		
D	PDB; 1K1M; 28-NOV-01.		
D	PDB; 1K1N; 28-NOV-01.		
D	PDB; 1K1O; 28-NOV-01.		
D	PDB; 1K1P; 28-NOV-01.		
D	PDB; 1MAX; 14-OCT-96.		
D	PDB; 1MAY; 14-OCT-96.		
D	PDB; 1MTS; 20-AUG-97.		
D	PDB; 1MTU; 12-NOV-97.		
D	PDB; 1MTV; 12-NOV-97.		
D	PDB; 1MTW; 12-NOV-97.		
D	PDB; 1N6X; 04-MAR-03.		
D	PDB; 1N6Y; 04-MAR-03.		
D	PDB; 1NTP; 15-APR-01.		
D	PDB; 1OYO; 29-APR-03.		
D	PDB; 1PPC; 31-JAN-94.		
D	PDB; 1PPE; 31-JAN-94.		
D	PDB; 1PPH; 31-JAN-94.		
D	PDB; 1QA0; 10-APR-00.		
D	PDB; 1QB1; 29-APR-00.		
D	PDB; 1QB6; 29-APR-00.		
D	PDB; 1QB9; 30-APR-00.		
D	PDB; 1OBN; 03-MAY-00.		
D	PDB; 1QBO; 03-MAY-00.		
D	PDB; 1QCP; 19-JAN-00.		
D	PDB; 1QL7; 25-AUG-00.		
D	PDB; 1QL8; 25-AUG-00.		
D	PDB; 1SBW; 07-APR-00.		
D	PDB; 1SFL; 09-JUL-99.		
D	PDB; 1SME; 31-JUL-94.		
D	PDB; 1TAE; 15-JUL-92.		
D	PDB; 1TAW; 24-JUN-97.		
D	PDB; 1TGB; 14-MAR-85.		
D	PDB; 1TGC; 09-APR-85.		
D	PDB; 1TGN; 22-OCT-84.		
D	PDB; 1TGS; 14-MAR-85.		
D	PDB; 1TGT; 15-OCT-90.		
D	PDB; 1TIO; 23-SEP-98.		
Query Match 17.5%; Score 394.5; DB 1; Length 243;			
Best Local Similarity 37.4%; Pred. No. 1e-24;			
Matches 91; Conservative 41; Mismatches 86; Indels 25; Gaps 8;			
QY	150 KIIGGEFTIENQWFAAIYRRHGGVYVYCGGSLISPCWVISATHCFIDYPKKEDYIV 209		
DB	20 KIVGGYTCGANTVPQVSL-----NSGYHFCGSLNSQWVSAHCY-----KSGIQV 68		
QY	210 YLGRSLNSQTQEMKFEVENLIHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQ 269		
DB	69 RLGEDNINVEGNEQFTISAKSIYHPSYNSNTL--NNDIMLIKLS---AASLSNRVAS 122		
QY	270 ICLPSMNDPQGTSCETIGFGKENSTDYLPQKMTVVVKLISHRECQPHYGVSEVTT 329		
DB	123 ISLPT--SCASAGTQCLISGWNTKSSGTSYDVLKCLKAPILSDSSCKSA--YFGQITS 178		
QY	330 KMLCAADPQWKTDCQGDGSGPLVCSLQGRMTLIGIVSWGRCALDKPGYTVRVSHFLP 389		
DB	179 NMFCAYLEGGKDCQSDSGGVVCS--GK--LQIVSWGSCAQKPKGYTVTKVNVYS 234		
QY	390 WIR 392		
DB	235 WIK 237		
RESULT 49			
TMS5 HUMAN STANDARD; PRT; 457 AA.			
ID	TMS5 HUMAN		
AC	Q9H3S3.		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Transmembrane protease, serine 5 (EC 3.4.21.-) (Spinesin).		
GN	TMPSRS5.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RA	PubMed=11741996;		
RT	Yamaguchi N, Okui A., Yamada T., Nakazato H., Mitsui S.;		
RT	"Spinesin/TMPSRS5, a novel transmembrane serine protease, cloned from		
RT	human spinal cord.";		
RL	J. Biol. Chem. 277:6806-6812(2002).		
CC	- - SUBCELLULAR LOCATION: Type II membrane protein (Potential).		
CC	- - TISSUE SPECIFICITY: Brain-specific. Predominantly expressed in		
CC	neurons, in their axons, and at the synapses of motoneurons in the		
CC	spinal cord.		
CC	- - SIMILARITY: Belongs to peptidase family S1.		
CC	- - SIMILARITY: Contains 1 SRCR domain.		

415 DTWRLVGVVSWGRACAEPNHGPVYAKVAEFLDWHDTAQD 454

DB

RESULT 50

TRY2_BOVIN	STANDARD;	PRT;	247 AA.
ID	Q29463;		
AC	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	trypsin, anionic precursor (EC 3.4.21.4).		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC	Bovidae; Bovinae; Bos.		
NCBI_TaxID=9913;			
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Holstein-Friesian; TISSUE=Pancreas;		
RA	MEDLINE=91065393; PubMed=1701147;		
RT	le Huerou I., Wicker C., Guilleto P., Toullec R., Puigserver A.;		
RT	"Isolation and nucleotide sequence of cDNA clone for bovine		
RT	pancreatic anionic trypsinogen. Structural identity within the		
RT	trypsin family.";		
RL	Eur. J. Biochem. 193;767-773(1990).		
CC	-1- CATALYTIC ACTIVITY: Preferential cleavage: ARG- -Xaa, Lys- -Xaa.		
CC	-1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).		
CC	-1- SUBCELLULAR LOCATION: Extracellular.		
CC	-1- SIMILARITY: Belongs to peptidase family S1.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; X54703; CAA38513.1; -		
DR	PIR; S13813; S13813.		
DR	HSP; P00763; IDPO.		
DR	MEROPS; S01.258; -		
DR	InterPro; IPR009003; Cys_Ser_trypsin.		
DR	InterPro; IPR001254; Peptidase_S1.		
DR	InterPro; IPR001314; Peptidase_S1A.		
DR	Pfam; PF00089; trypsin; 1.		
DR	PRINTS; PR00722; CHYMOTRYPSIN.		
DR	SMART; SM00020; TRYP_SPC; 1.		
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.		
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.		
DR	PROSITE; PS00135; TRYPSIN_SER; 1.		
DR	Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;		
KW	Calcium-Binding; Signal.		
FT	FT SIGNAL 1 15		POTENTIAL.
FT	FT PROPEP 16 23		ACTIVATION PEPTIDE.
FT	FT CHAIN 24 247		TRYPSIN, ANIONIC.
FT	FT ACT_SITE 63 63		CHARGE RELAY SYSTEM.
FT	FT METAL 75 75		CALCIUM (BY SIMILARITY).
FT	FT METAL 77 77		CALCIUM (VIA CARBONYL OXYGEN)
FT			(BY SIMILARITY)
FT	FT METAL 80 80		CALCIUM (VIA CARBONYL OXYGEN)
FT			(BY SIMILARITY)
FT	FT METAL 85 85		CALCIUM (BY SIMILARITY).
FT	FT ACT_SITE 107 107		CHARGE RELAY SYSTEM.
FT	FT ACT_SITE 200 200		CHARGE RELAY SYSTEM.
FT	FT DISULFID 30 160		BY SIMILARITY.
FT	FT DISULFID 48 64		BY SIMILARITY.
FT	FT DISULFID 132 233		BY SIMILARITY.
FT	FT DISULFID 139 206		BY SIMILARITY.
FT	FT DISULFID 171 185		BY SIMILARITY.
FT	FT DISULFID 196 220		BY SIMILARITY.
FT	FT SITE 194 194		REQUIRED FOR SPECIFICITY.
FT	SEQUENCE 247 AA; 26288 MW; 50A070495A7731DB CRC64;		

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

O protein - protein search, using sw model
R on: May 25, 2004, 14:44:05 ; Search time 66.7546 Seconds
(without alignments)
1904.795 Million cell updates/sec

T ie: US-09-880-503-6
P ect score: 2257
S tence: 1 SNELHQVSNCLNGTCTV.....VSHFLPWIRSHKENGIAL 403

S ring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

S ched: 1017041 seqs, 315518202 residues

T al number of hits satisfying chosen parameters: 1017041

M lnum DB seq length: 0
M lnum DB seq length: 2000000000

P -processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

D abase :
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

R	alt	Score	Query	Match	Length	DB	ID	Description
1	1885	83.5	433	6	Q8MIL0	oryctolagus		
2	1880	83.3	433	6	Q8MIL0	oryctolagus		
3	953.5	42.2	214	6	Q8MIL0	oryctolagus		
4	858.5	38.0	562	4	Q86YK8	homo sapien		
5	852.5	37.8	516	4	Q8MIL0	oryctolagus		
6	823	36.5	564	4	Q8MIL0	oryctolagus		
7	815.5	36.1	562	6	Q8S223	sus scrofa		
8	808.5	35.8	231	11	Q8C6L2	mus musculus		
9	788.5	34.9	395	4	Q9BZW1	homo sapien		
10	778	34.5	154	4	Q96SE8	homo sapien		
11	722.5	32.0	154	13	Q800Y7	meleagris g		
12	715	31.7	653	11	Q8VCS4	mus musculus		
13	680.5	30.2	616	6	Q97507	sus scrofa		
14	678	30.0	615	4	Q81Z25	homo sapien		
15	668	29.6	597	11	Q35727	mus musculus		
16	668	29.6	609	11	Q80YC5	mus musculus		

Q8X0D2 mus musculus
Q14520 homo sapien
Q9TVAS bos taurus
Q97587 oryctolagus
Q9R0W3 rattus norv
Q15146 homo sapien
Q95M89 equus cabal
Q96AF3 homo sapien
Q46506 papio hamad
Q8AVB0 brachydania
Q18783 macropus eu
Q9Y1V3 polyanthroca
Q8N171 homo sapien
Q8NF86 homo sapien
Q866V4 halocynthia
Q86YM4 homo sapien
Q9BYE1 homo sapien
Q99J88 rattus norv
Q9BK47 ludia foli
Q9BYE2 homo sapien
Q8CFE0 mus musculus
Q80WM7 mus musculus
Q804X7 gallus gall
Q8CGR6 mus musculus
Q8R0P5 mus musculus
Q8CJ16 rattus norv
Q8CJ17 rattus norv
Q8CJ17 mus musculus
Q8WV21 homo sapien
Q9DAT3 mus musculus
Q7Z410 homo sapien
Q7Z411 homo sapien
Q9J117 rattus norv
Q91V47 mus musculus
Q86899 scolopendra
Q8TTR0 canis famil
Q7Z5F4 homo sapien
Q812A6 mus musculus
Q91444 mus musculus
Q7SV86 xenopus lae
Q8BIK6 mus musculus
Q8BIV6 mus musculus
Q8QV0 cyprinus ca
Q9DC86 mus musculus
Q9CR35 mus musculus
Q7ZT70 lampetra ja
Q9DGR2 xenopus lae
Q9QV4 mus musculus
Q9QV3 m hippotas
Q8101 herdmantia m
Q9DGR1 xenopus lae
Q8N2U3 homo sapien
Q90865 gallus gall
Q81UW0 homo sapien
Q96R28 homo sapien
Q9NJS5 anopheles g
Q9BH03 felis silve
Q867B7 canis famil

ALIGNMENTS

RESULT 1
Q8MIL0
ID Q8MIL0; PRELIMINARY; PRT; 433 AA.
AC Q8MIL0;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator.
GN PLAU.

Oryctolagus cuniculus (Rabbit).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NCBI_TaxID=9986;
[1]
SEQUENCE FROM N.A.
MEDLINE=2215945; PubMed=12149463;
Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
Dichek D.A.;
"Increased expression of urokinase during atherosclerotic lesion
development causes arterial constriction and lumen loss, and
accelerates lesion growth";
Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
EMBL; AY122285; AAM83187.1; -.
GO; GO:0004263; F:chymotrypsin activity; IEA.
GO; GO:0016301; F:kinase activity; IEA.
GO; GO:0008233; F:peptidase activity; IEA.
GO; GO:0004235; F:trypsin activity; IEA.
GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
InterPro; IPR009003; Cys_Ser_trypsin.
InterPro; IPR006209; EGF-like.
InterPro; IPR000001; Kringle.
InterPro; IPR001254; Peptidase S1.
InterPro; IPR001314; Peptidase_S1A.
InterPro; IPR008293; Pept_S1A_UPA.
Pfam; PF00051; Kringle; 1.
Pfam; PF00089; trypsin; 1.
PROSITE; PS00022; EGF_1; 1.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS0070; KRINGLE_2; 1.
PROSITE; PS0240; TRYPSIN_DOM; 1.
PROSITE; PS0134; TRYPSIN_HIS; 1.
PROSITE; PS0135; TRYPSIN_SER; 1.
PIRSP; PIRSF001144; Urk_plasm_act; 1.
Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SEQUENCE 433 AA; 48375 MW; 65E64F364154980 CRC64;
Query Match 83.5%; Score 1885; DB 6; Length 433;
Best Local Similarity 82.8%; Pred. No. 9.5e-170;
Matches 342; Conservative 23; Mismatches 38; Indels 10; Gaps 2;
1 SNEHQV--PNCDCNGTCTVSNKYPSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFY 58
21 SHELHGVSDASNCGLNGTCTVYKIFSNWRCNCPKFKQGEHCCEIDTLKTCYHGDHSY 80
59 RGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQV 118
81 RKGANTDMDRPLCLAWNSANVLTNTYHAHRPDALQLGLGKHNYCRNPDHQRPPWCYVQV 140
119 LKPLVQECVHDCAD-----GKLKFCQCGKTLRPRFKLIGGEFTTIENQWFAAIYR 170
141 LKLIQECVHDCSSGKKPALPPKLEFQCGQKALRPRFKLIGGEFTTIENQWFAAIYR 200
171 RHRGGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVVILGRSLNNTQGMKPEVEN 230
201 RHRGGSVTVYCGSLISPCWVVSATHCFINQKEDYIVVILGRSLNMTPEGKPEVEQ 260
231 LILKVDYADTLAHNDIALKIRSGRCHQPSRTTQTCLPSMYNDPFGTSCBITGF 290
261 LILHEGYSADTLAHNDIALKILSNNGQCAQPSRSITCLLPWNADPNFGTSCBITGF 320
291 GKENSTLYLPEQLKMTVKLISHRECOQPHYGSVTTKMLCAADPQWKTDCQGDSSG 350
321 GKENSTLYLPEQLKMTVKLVSVQECQPHYGSVTTKMLCAADPQWKTDCQGDSSG 380
351 PLVCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHFLPWIRSHTKENGIAL 403

Db 381 PLVCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHFLPWIRSHTKENGIAL 433
RESULT 2
Q8MHY7
ID Q8MHY7 PRELIMINARY; PRT; 433 AA.
AC Q8MHY7;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the ENBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT complete cds";
RL Submitted (JUL-2002) to the ENBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029517; AAK40239.1; -.
DR EMBL; AB087224; BAC02685.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004235; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR008293; Pept_S1A_UPA.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS0070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; 1.
DR PROSITE; PS0135; TRYPSIN_SER; 1.
DR PIRSF; PIRSF001144; Urk_plasm_act; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SEQUENCE 433 AA; 48444 MW; 6DD3A5A71010A6EE CRC64;
Query Match 83.3%; Score 1880; DB 6; Length 433;
Best Local Similarity 82.6%; Pred. No. 2.8e-169;
Matches 341; Conservative 23; Mismatches 39; Indels 10; Gaps 2;
QY 1 SNEHQV--PNCDCNGTCTVSNKYPSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFY 58
DB 21 SHELHGVSDASNCGLNGTCTVYKIFSNWRCNCPKFKQGEHCCEIDTLKTCYHGDHSY 80
QY 59 RGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQV 118
DB 81 RKGANTDMDRPLCLAWNSANVLTNTYHAHRPDALQLGLGKHNYCRNPDHQRPPWCYVQV 140
QY 119 LKPLVQECVHDCAD-----GKLKFCQCGKTLRPRFKLIGGEFTTIENQWFAAIYR 170


```

D 141 LKQLTQCKVHDCSSGKKPALPPGKLEFQCCKQKALRPRFKLIGGFTTIENQPFPAAIYR 200
Q 171 RRRGGSVTVVCGSLSPCWISATHCFIDYFKKEDYIVYLGSRSLNSNTGEMKFEVEN 230
D 201 RRRGGSVTVVCGSLSPCWVVSATHCFINQKKEDYIVYLGSRSLNSMTPEGNKFEVEQ 260
Q 231 LILHKDYSADTLAHNDIALLKIRSEKRCQAPSTIOTICLPSMYNDPQFTSCETGTF 290
D 261 LILHGGYRADTLAHNDIALLKILSNNGCQAPSRSIOTICLPPWNADEFGTSCETGTF 320
Q 291 GKENSTDYLYPQLKNTVVKVLSHRECCQPHYGVSEVTTKMLCAADPQWKTDCQGDGG 350
D 321 GKENSTDYLYPQLKNTVVKVLSVQCCQPHYGVSEVTTKMLCAADPQWKTDCQGDGG 380
Q 351 PLVCSLOHMTLTGVSNGRCALKDKGVTVTRVSHFLPWIRSHRTKEENGLAL 403
D 381 PLVCSVQGRMTLTGIVSMRGCKLNKPGVTVTRVSRFLPWIRSHRTKEENGLAL 433

E ULT 3
Q 99X770 PRELIMINARY; PRT; 214 AA.
A 01-NOV-1999 (T-EMBLrel. 12, Created)
D 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
D 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
D Urokinase-type plasminogen activator (Fragment).
C Oryctolagus cuniculus (Rabbit).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
C NCBI_TaxID=9986;
K [1]
K SEQUENCE FROM N.A.
K TISSUE=Lung;
K Vin J., Idell S.,
K "Partial mRNA of rabbit uPA.";
K Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
K -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
C -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
C EMBL; AF097647; AAD39351.1; -.
D HSP; P00749; 1EJN.
D MEROS; S01.231; -.
D GO; GO:0004263; F:chymotrypsin activity; IEA.
D GO; GO:0016301; F:kinase activity; IEA.
D GO; GO:0008233; F:peptidase activity; IEA.
D GO; GO:0004295; F:trypsin activity; IEA.
D GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
D InterPro; IPR009003; Cys Ser trypsin.
D InterPro; IPR000001; Kringle.
D InterPro; IPR001254; Peptidase_S1.
D InterPro; IPR001314; Peptidase_S1A.
D Pfam; PF00051; kringle; 1.
D Pfam; PF00089; trypsin; 1.
D PRINTS; PR00722; CHYMOTRYPSIN.
D ProDom; PD000395; Kringle; 1.
D SMART; SM00130; KR; 1.
D SMART; SM00020; Tryp_SPC; 1.
D PROSITE; PS00022; EGF_1; 1.
D PROSITE; PS01186; EGF_2; 1.
D PROSITE; PS01253; FIBRONECTIN_1; 1.
D PROSITE; PS00021; KRINGLE_1; 2.
D PROSITE; PS00070; KRINGLE_2; 2.
D PROSITE; PS50240; TRYPSIN_DOM; 1.
D PROSITE; PS00134; TRYPSIN_HIS; 1.
D PROSITE; PS00135; TRYPSIN_SER; 1.
D SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match 42.2%; Score 953.5; DB 6; Length 214;
Best Local Similarity 81.3%; Pred. No. 5.5e-82;
Matches 174; Conservative 12; Mismatches 21; Indels 7; Gaps 1;

Q 67 MGRPCLPWNATVLOOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLKPLVQEC 126
D 1 MDRPCLAWNANVLTQTYHAHRPDALQLGLGKHNYCRNPNDRPWCYVQVGLKQLIQEC 60

```

```

QY 127 MYHDC-----DGKLFQCCGKTLRPRFKLIGGFTTIENQPFPAAIYRHRGGSVTV 179
DB 61 KVHDSGKKPALPPGKLEFQCCKQKALRPRFKLIGGFTTIENQPFPAAIYRHRGGSVTV 120
QY 180 VCGSLSPCWISATHCFIDYFKKEDYIVYLGSRSLNSNTGEMKFEVENLILHKDYS 239
DB 121 VCGSLSPCWVVSATHCFINQKKEDYIVYLGSRSLNSMTPEGNKFEVEQLILHEGYRA 180
QY 240 DTLAHNDIALLKIRSEKRCQAPSTIOTICLIP 273
DB 181 DTLAHNDIALLKILSNNGCQAPSRSIOTICLIP 214

RESULT 4
Q66YK8 PRELIMINARY; PRT; 562 AA.
ID Q66YK8
AC Q66YK8; 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Tissue plasminogen activator.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu Y., Xu L., Zeng Y., He X.;
RT "cDNA of tissue plasminogen activator";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY221101; AAC34406.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006203; EGF_like.
DR InterPro; IPR000083; Fibrinctn.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
D SQ SEQUENCE 562 AA; 62902 MW; 837D98392F6EDD1F CRC64;

Query Match 38.0%; Score 858.5; DB 4; Length 562;
Best Local Similarity 37.5%; Pred. No. 1.8e-72;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

```

```

QY 3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWNCPPKFGQCHCEIDKSKTCYEGNGHFY 58
DB 77 QCHSVVPKSCSEPRFCFGGTCQALYFSDP-VCCQPEGFAGKCCBIDTRATCYEDQGISY 135
QY 59 RGKASTDTMGPRCLPWNATVLOOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVQ 118
DB 136 RGTWSTAESGECTNWNSSALAAQKPYSGRRPDALRLGLGNHNYCRNPNDRSKPCWYVFK 195

```

QY 119 LKPLVQECMVHDCADG----- 134
 DE 196 GYSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNMILIGNVYTAQNS 255
 QY 135 -----KLKQ-----CG-QKTLRPRFKLIGGE 155
 DE 256 AALGLGKHNCRNPDGDAKPCWCHLVKNRRLTWEYCDVPSCSTCGLRQYSQFRIKGL 315
 QY 156 FTTIENQWPAAYRRH-RGGSVTVYCGSLISPCWVISATHCFIDYPPKEDYVIVLGRS 214
 DE 316 FADIASHPWOAAIPAKHRRSGRPFGLCGGILISSCWLSAAHCFQERFPPHLLTVILGRT 375
 QY 215 RLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPS 274
 DE 376 YRVVPGEBEOKFEVEKIVHKEFDDDT--YNDIALQLKSDSSRCAQESSVVRTVCLPP 433
 QY 275 MYNDPQGTSCBITGFGKNSDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCA 334
 DE 434 ADLQLPDWTCELSGKGHEALPFYSERLKEAHLVLPSSRCTSQHLLNRTVTDNMLCA 493
 QY 335 AD-----PQWKT--DSCQDGGPLVCSLOGRMVLTGIVSGRGCAKDKPGVTVRSHL 388
 DE 494 GDTSGGPOANLHDACQDGGPLVCLNDGRMTLVGLIISWGLGCGQKDPGVTYKRVNYL 553
 QY 389 PWIRSHTK 396
 DE 554 DWIRDNR 561
 RE LT 5
 QY 99 Q9BU99 PRELIMINARY; PRT; 516 AA.
 ID Q9BU99; AC
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to plasminogen activator, tissue.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DE HSP; P00750; 1A5H.
 DE GO; GO:0004253; F:chymotrypsin activity; IEA.
 DE GO; GO:0008233; F:peptidase activity; IEA.
 DE GO; GO:0004295; F:trypsin activity; IEA.
 DE GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DE InterPro; IPR003003; Cys_Ser_trypsin.
 DE InterPro; IPR006209; EGF-like.
 DE InterPro; IPR006210; IEGF.
 DE InterPro; IPR000001; Kringle.
 DE InterPro; IPR001254; Peptidase_S1.
 DE InterPro; IPR001314; Peptidase_S1A.
 DE Pfam; PF00008; EGF; 1.
 DE Pfam; PF00051; kringle; 2.
 DE Pfam; PF00089; trypsin; 1.
 DE PRINTS; PR00722; CHYMOTRYPSIN.
 DE PRINTS; PR00018; KRINGLE.
 DE PRODOM; PD000395; Kringle; 2.
 DE SMART; SM00181; EGF; 1.
 DE SMART; SM00130; KR; 2.
 DE SMART; SM00020; Tryp_Spc; 1.
 DE PROSITE; PS00022; EGF_1; 1.
 DE PROSITE; PS01186; EGF_2; 1.
 DE PROSITE; PS00021; KRINGLE_1; 2.
 DE PROSITE; PS00070; KRINGLE_2; 2.

DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;
 Query Match 37.8%; Score 852.5; DB 4; Length 516;
 Best Local Similarity 37.1%; Pred. No. 6e-72;
 Matches 184; Conservative 55; Mismatches 154; Indels 103; Gaps 9;
 QY 1 SNEHL-----QVPSNCDCLNGGTCSNVKYSFIHNCNCPKFGGHCCEIDKSKTC 50
 DE 23 SETHAFRRGARYQCCSEPCFNGGTCCQALYFSDF-VCCDEGAGKCEIDTRATC 81
 QY 51 YEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNCRNPNRR 110
 DE 82 YEDQGISYRGTWSTAESGAECTNNWSSALAQKPYSGRRPDAILRLGLGNHNYCRNPD 141
 QY 111 PWCYVQVGLKPLVQECMVHDCADG----- 134
 DE 142 PWCYVFKAGKYSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNMILIGK 201
 QY 135 -----KLKQ-----CG-QKTLRP 147
 DE 202 VYTAQNPSAQAALGLGKHNCRNPDGDAKPCWCHLVKNRRLTWEYCDVPSCSTCGLRQYSOP 261
 QY 148 RPKIIGGEFTTIENQWPAAYRRH-RGGSVTVYCGSLISPCWVISATHCFIDYPPKED 206
 DE 262 QFRINGGLFADIASHPWOAAIPAKHRRSGRPFGLCGGILISSCWLSAAHCFQERFPPHH 321
 QY 207 YIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKRGCAQPSRT 266
 DE 322 LTVILGRYRVVPGEBEOKFEVEKIVHKEFDDDT--YNDIALQLKSDSSRCAQESSV 379
 QY 267 ICTCLPSMYNDPQGTSCBITGFGKNSDYLYPEQLKMTVVKLISHRECCQPHYGVSE 326
 DE 380 VRTVCLPPADLQLPDWTCELSGKGHEALPFYSERLKEAHLVLPSSRCTSQHLLNRT 439
 QY 327 VTTKMLCAAD-----PQWKT--DSCQDGGPLVCSLOGRMVLTGIVSGRGCAKDKPGV 380
 DE 440 VTDNMLCAGDTRSGGPOANLHDACQDGGPLVCLNDGRMTLVGLIISWGLGCGQKDPGV 499
 QY 381 YRVVSHFLPWIRSHTK 396
 DE 500 YTKVTNYLDWIRDNR 515
 RESULT 6
 Q8MKB1 PRELIMINARY; PRT; 564 AA.
 ID Q8MKB1; AC
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Tissue-type plasminogen activator.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DE EMBL; AY029518; AAK40240.1; -.
 DE GO; GO:0005576; C:extracellular; IEA.
 DE GO; GO:0004263; F:chymotrypsin activity; IEA.
 DE GO; GO:0008233; F:peptidase activity; IEA.
 DE GO; GO:0004295; F:trypsin activity; IEA.
 DE GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DE InterPro; IPR009003; Cys_Ser_trypsin.

D InterPro; IPR006209; EGF like.
D InterPro; IPR000083; Fibrinctnl.
D InterPro; IPR006210; IEGF.
D InterPro; IPR000001; Kringle.
D InterPro; IPR001254; Peptidase S1.
D InterPro; IPR001314; Peptidase_S1A.
D Pfam; PF00008; EGF; 1.
D Pfam; PF00039; fni; 1.
D Pfam; PF00051; kringle; 2.
D Pfam; PF00053; trypsin; 1.
D PRINTS; PR00722; CHYMOTRYPSIN.
D PRINTS; PR00018; KRINGLE.
D ProDom; PD000395; Kringle; 2.
D SMART; SM00181; EGF; 1.
D SMART; SM00058; fni; 1.
D SMART; SM00130; KR; 2.
D SMART; SM00020; Tryp_Spc; 1.
D PROSITE; PS00022; EGF_1; 1.
D PROSITE; PS01186; EGF_2; 1.
D PROSITE; PS01253; FIBRONECTIN 1; 1.
D PROSITE; PS00021; KRINGLE 1; 2.
D PROSITE; PS00070; KRINGLE 2; 2.
D PROSITE; PS00240; TRYPSIN DOM; 1.
D PROSITE; PS00134; TRYPSIN HIS; 1.
D PROSITE; PS00135; TRYPSIN_SER; 1.
D EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
D Serine protease.
E SEQUENCE 564 AA; 62726 MW; 459D8BAC6D4A937C CRC64;
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
Query Match 36.5%; Score 823; DB 6; Length 564;
Best Local Similarity 36.3%; Pred. No. 4.1e-69;
Matches 178; Conservative 55; Mismatches 148; Indels 110; Gaps 9;
C S: HQVP-SNCD---CLNGGTCVSNKFNHWCNPKKFGQGHCEID-----45
D 80 HSPVQSCSPRCLNGGTCQALYFSDP-VCQCPGVGKRCBVDPRACYEDRGIGYRG 138
C 46:-----45
D 139 TWSTTESGAOCVNNSSWLALKPYSGRKNALRLGLGNHNYCNPNRDRTPKCYVFRAGT 198
C 46:-----KSKTCYEGNGHGYRGRKASTDTMGRCPLPNSATVLQOYTHAHRSDA 91
I 199 YSPFCSTPACSKKNGCNVYLGKQAVRGTHSLTSSGASCLPNSMLLVGEKVTARQNA 258
C 92 LQGLGKHNYCRPNDRRRWCYVQVGLKPLVCECMWHDCAQKLFQCG-QKTLPAPRK 150
I 259 EALGLGKHNYCRPNDRGDSKPCWCHLVKNRKLVIYCDVPQCA-----TCGLRQDKQPPR 312
C 151 IIGGEFTTIEQWFAAIY-RRHRGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIV 209
D 313 IKGLFTDITAHQWQAIFTNERSPGERFLOGGILLNSCWLSAAHCFLEPPQOKLRY 372
C 210 YLGRSLNNTQGEKMFVENLILHKDYSADTLAHNDIALLKIRSGRCQAQPSRTIQT 269
D 373 ILGRTYPLVSAEAEQIFVEQPIHERFDEGT--YNDIALLLKLTSGSCAQESQAVRL 430
C 270 ICLPSMYNDPQFGTSCEITGFKENSTDYLYPEQLXMTVVKLISHRECQPHYSGEVT 329
D 431 VCLPDASLQLPDWTCELSYGVKHEEFPVSFQLKXAHVRLVPSRCLTPQQLKRTVTG 490
C 330 KMLCAADPQW-----KTDSCQSDGGPLVCSLQGRMTLTGIVSWRGKCALDKPGVYTR 383
D 491 NMLCAGTRSGGAQVNLHDACQSDGGPLVCMTDGHTLTIGIISWGLGCGQKDVPGVYTK 550
C 384 VSHFLPWIRSH 394
D 551 VVNYLGIQOH 561

AC Q8SQ23;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Enamel organ;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBSJ databases.
CC 1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AAC00297.1; -;
DR HSBP; P00761; IANI.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR008209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fni; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00053; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; fni; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 1.
DR PROSITE; PS00021; KRINGLE 1; 2.
DR PROSITE; PS00070; KRINGLE 2; 2.
DR PROSITE; PS00240; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101E8 CRC64;

Query Match 36.1%; Score 815.5; DB 6; Length 562;
Best Local Similarity 35.9%; Pred. No. 2.1e-68;
Matches 175; Conservative 60; Mismatches 156; Indels 97; Gaps 11;
QY 3 ELHQP-SNCD---CLNGGTCVSNKFNHWCNPKKFGQGHCEIDSKSKTCYEGNGHGY 58
DB 77 QCHSVPKSGEPKCFNGTCLQAIYFSDP-VCQCPGVGFRQCEIDARATCYEDQGIT 135
QY 59 RGKASTDTMGRCPLPNSATVLQOYTHAHRSDALQGLGKHNYCRPNDRRRWCYVQVG 118
DB 136 RGTWSTTESGAECVNNNTSGLASMPYNGRRPDAVKLGLGNHNYCRNPKDKSKPCYIFKA 195
QY 119 LK-----PLV-----QECNV-----128
DB 196 EKYSPDFCSTACTKEKEECYTGKLDYRGTRSTMSGAFCLPNSLVLMGKIITANNNS 255
QY 129 -----HDCADGKLKFO-----CGQKTLR-PRFKIIGGE 155
DB 256 AQTGLGKHNYCRNPDPDGTQPCWCHLVKDHKLITWEYCDLPQCVTCGLAQYKEPQRIKGL 315

QY 156 FTTIENQWPAALY-RRHGGSVYVCGGSLISPCWVISATHCFIDYPPKEDYIVVLGRS 214
DB 316 YADITSHFWAAIFVKNRSGPGRFLCGGLISSCWLSAAHCFQBRPPHVRVVLGRT 375
QY 215 RLNSNTQGENKFEVENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIQTCLPS 274
DB 376 YRLVPGSEEAFAVEKYIVHKEFDDDT--YDNDIALQLKSDSLTCAQSSDAVRTVCLPE 433
QY 275 MYNDPQGTSCBTGTGKENSITDYLPEQLKMTVVKLISHRECOQPHYYGSEVTTMLCA 334
DB 434 ANQLPDPWTECELSGVGKHSAPFFYSERLUKEAHRVLPYSSRCTSKHLFNKTIINMLCA 493
QY 335 ADPOW-----KTDSCGDSGGPLVCSLQGRMTLTGIVSNGRCALKDKPGVYVTRYSHEL 388
DB 494 GDRSGGDNANLHADCGDSGGPLVCMKGNHMTLVGVISNGLGCGQKDPGVYVTKVTNVL 553
QY 389 PWIRSHTK 396
DL 554 NWIRDNTR 561
RE 57 8
L2 Q86L2 PRELIMINARY; PRT; 231 AA.
AC Q86L2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RF STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12456851;
RA The RANOM Consortium,
RA the RANOM Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR EMBL; AK054349; BAC35743.1; -.
DR PIR; P0534; P0534.
DR GO; GO:0004295; P:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00083; trypsin; 1.
DR PRINTS; PR00083; trypsin; 1.
DR Pfam; PF00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR NON TER 231
FT SEQUENCE 231 AA; 25510 MW; 25E8980A682737F2 CRC64;

QY 129 HDCADGKLGK-----FQCGQKTLRPRFKIIGGEFTTIENQWPAALYRRHGGSV-VTY 179
DB 150 HDCSLSKKPSVVDQGFQCGQKALRPRFKLVGGFTEVENQWPAALYQKNGGSPSP 209
QY 180 VCGGSLISPCWVISATHCFI 199
DB 210 KCGGSLISPCWVASAAHCFI 229
RESULT 9
Q8BZWI
ID Q8BZWI PRELIMINARY; PRT; 395 AA.
AC Q8BZWI;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neonatal thrombolytic agent alpha-form (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF260825; AAK11956.1; -.
DR HSSP; P00750; 1PK2.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004283; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON TER 395
FT SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 34.9%; Score 788.5; DB 4; Length 395;
Best Local Similarity 42.9%; Pred. No. 4.9e-66;
Matches 161; Conservative 49; Mismatches 144; Indels 21; Gaps 8;
QY 30 WCNCPKFGGQHCIDKSKTCYEGNGHFGYRGKASTDTMGRECLPWNSTVLCQTYHAHRS 89
DB 33 WCNLS---GRAQCS-EGNSDCYFGNGSAYRGTHSLTSGASCLPWNMSILIGKVYTAQN 87
QY 90 DALQLGLGHKYNCRNPNRPPRCVQVGLPVLQECMVHDCADGKLKFCQG-QKTLRPR 148
DB 88 SAQALGLGHKYNCRNPDGDAKFWCHLVKNRRLTWECYCDVPSCS-----TCGLQKYSOPQ 141
QY 149 FKIIIGGEFTTIENQWPAALYRRH-RGGSVTVVCGGSLISPCWVISATHCFIDYPPKEDY 207

Query Match 35.8%; Score 808.5; DB 11; Length 231;
Best Local Similarity 69.5%; Pred. No. 3.2e-68;
Matches 139; Conservative 21; Mismatches 31; Indels 9; Gaps 2;
QY 9 SNCDLGGTCVSNKYFSNIHWNCNPKFGQHCIDKSKTCYEGNGHFGYRGKASTDTMG 68
DB 30 SNCCQNGGVCSYKYFSNIRCRSCPRKFGQHCIDKSKTCYEGNGHFGYRGKASTDTMG 89
QY 69 RPLPWNSTVLCQTYHAHRS DALQLGLGHKYNCRNPNRPPRCVQVGLKPLVQECMV 128
DB 90 RPLAWNAPAVLQKPYNAHRPDALSLGLGHKYNCRNPDNQKRWPCYVQIGLRQVQECMV 149

D 142 FRKGLFLADIASHFWQAAIFAKHEESPOERFLOGGILISSCWILSAAHCFORPPPHL 201
Q 208 IYVIGSRSLNSTQGMKEFEVENLILHKDYSADTLAHNDIALKIRSKGRCQAPSRTI 267
D 202 TWLGTIRVWPGEEBQCEKEYIVHKEFDDDT--YNDIALQLKSDSECAESSVW 259
Q 268 QITCLPSMNDPOFGTSCITGFKENSTDYLYPQLKTVVYLISHRECQOPHYGSEV 327
D 260 RTVCLPFPADQLPDMTECELSGKGREALSPYSEKLEAHVRLYPSSRCTSQHLLNETV 319
Q 328 TTMQLCAAD-----POWKT-DSQGDSDGGLVCSLQGRMTLTGIVSGWGRGCALKDKPGVY 381
D 320 TDNMLCAGTRSGGPPANLHDACQDGGPLVCLNDGRMTLVGLISWGLGCGQKQVPGVY 379
Q 382 TRVSHLPWIRSHTK 396
D 380 TKVTNYLDWIRDMR 394

R JUL 10
C SE8 PRELIMINARY; PRT; 154 AA.
Q96S88;
D 01-DRC-2001 (TrEMBLrel. 19, Created)
D 01-DRC-2001 (TrEMBLrel. 19, Last sequence update)
D 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
D Urokinase-type plasminogen activator amino-terminal fragment.
C XFF.
C Homo sapiens (Human).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
C NCBI_TaxID=9606;
R E SE8 SEQUENCE FROM N.A.
R Fu J., Bai X., Ruan C.;
R "Cloning and expression of the amino-terminal fragment of human
R urokinase-type plasminogen activator."
R Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
R [2]
R SEQUENCE FROM N.A.
R Bai X., Fu J., Wang W., Xi X., Ruan C.;
R "Overexpression of the amino-terminal fragment of human urokinase-type
R plasminogen activator in breast cancer cells results in decreased
R tumor invasion, growth and angiogenesis."
R Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
R -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
D EMBL; AV029537; AAK38734.1; --
D GO: GO:0016301; P:kinase activity; IEA.
D InterPro; IPR006210; EGF.
D InterPro; IPR006209; EGF-like.
D Pfam; PF00051; kringle; 1.
D PRINTS; PR00018; KRINGLE.
D ProDom; PD000395; Kringle; 1.
D SMART; SM00181; EGF; 1.
D SMART; SM00130; KR; 1.
D PROSITE; PS00022; EGF_1; 1.
D PROSITE; PS00021; KRINGLE_1; 1.
D PROSITE; PS00070; KRINGLE_2; 1.
D Glycoprotein; Kinase; Kringle.
R SEQUENCE 154 AA; 17305 MW; A3CCF2FCFF505572 CRC64;

Query Match 34.5%; Score 778; DB 4; Length 154;
Best Local Similarity 99.3%; Pred. No. 1.5e-85;
atches 133; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

C 1 SNELHVPNSCDCLNGGTCVSNKYFNSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFRG 60
D 21 SNELHVPNSCDCLNGGTCVSNKYFNSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFRG 80
C 61 KASTDTMGRCLPWNATVLQOTYHAHRSDALQGLGKNYCNPNRPRPCVYQVGLK 120
D 61 KASTDTMGRCLPWNATVLQOTYHAHRSDALQGLGKNYCNPNRPRPCVYQVGLK 140

QY 121 PLVQECMVHDCADG 134
DB 141 LLVQECMVHDCADG 154

RESULT 11
Q800Y7
ID Q800Y7 PRELIMINARY; PRT; 540 AA.
AC Q800Y7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hepatocyte growth factor activator (Fragment).
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2212796; PubMed=12128063;
RA Holsberger D.R., Becker A.E., Thurston R.J., Rice C.D.;
RT "Expression of a hepatocyte growth-factor activator protein in turkey
RT (Meleagris gallopavo) deferent duct epithelial cells."
RL Comp. Biochem. Physiol. 132:769-777(2002).
[2]
RN
RP SEQUENCE FROM N.A.
RA Holsberger D.R., Becker A.E., Thurston R.J., Rice C.D.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY216598; AAO46038.1; --
DR GO: GO:0005576; C:extracellular; IEA.
DR GO: GO:0004263; F:chymotrypsin activity; IEA.
DR GO: GO:0004295; F:trypsin activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_crypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; TBGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PS00722; CHYMOTRYPSIN.
DR PRINTS; PS00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
FT NON_TER 1
FT NON_TER 540
SQ SEQUENCE 540 AA; 61159 MW; 0BB3B4A89C0B577F CRC64;

Query Match 32.0%; Score 722.5; DB 13; Length 540;
Best Local Similarity 38.3%; Pred. No. 1.3e-59;
Matches 155; Conservative 60; Mismatches 147; Indels 43; Gaps 11;

QY 13 CLNGGTC-----VSNKYFNSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFRGKASTDTMG 68
DB 143 CMNGGECMIASSGKTV-----CDCKGPFVKGKCNVFNHHCYRGNGTEYRGTAKTISG 197

QY 69 RPCLPWN SATVLQOTVHAHRS DALQLGLGKHNYCNPDRRRPWCYVQVGLKELVQECMV 128
 DB 198 HSCLPWN SDLLYRELHVDSEKAVOLGGLGFPSCRNPDDEDEKFWCTIMKDNLSWYECNI 257
 QY 129 HDCAADGKL-----KFCQCK-----TLRPFKIIGBFTTIENQWPFAAY 169
 DB 258 TSCASRRRPPVLEIDITFAVPRRCGRHKKGSFVRPR--IIGSSSLPGSHPTAAIV 315
 QY 170 RRHGGGVTVCGSSLSPCWVISAHCFTDYPKEDYIVYLGSRNSNTQCEMKFEVE 229
 DB 316 ---IGES---FCAGTLLOTWVSAAHCFANSPQKSIKVVLCQHFNRTTDTVQTFEIE 369
 QY 230 NLILHKDYSADTLAHNDIALKIRSKEGRCAPQRTIQTICLP---SMYNDPQFTSC 286
 DB 370 KYILYPOYSVFRPREH-DIALIKLKNQGRCAVKSQVQICLPESNTVFPD-QF--KCQ 425
 QY 287 ITGSGKENS DYLPEOLKMTVVKLISHRECQPHYVGSVTTKMLCAADPQWKTSCOG 346
 DB 426 ISGWHKHENITGSDVLQETLPIIPEEKCRGPEIYGTETSENNFCAGYFDSKSDACOG 485
 QY 347 DSGGLVCSLQGRMTLTGIVSWGRGCALDKPGVYTRVSHFLPWI 391
 DB 486 DSGGLACENNEISYLGVISWGDGCRVKNKPGVYTRVSNVNI 530

RE 12
 ID 34 Q8VCS4 PRELIMINARY; PRT; 653 AA.
 AC Q8VCS4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RE SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Strauberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC GO:0004295; F:trypsin activity; IEA.
 DE HSSP; P00761; 1AN1.
 DE GO:0005576; C:extracellular; IEA.
 DE GO:0004283; F:chymotrypsin activity; IEA.
 DE GO:0008233; F:peptidase activity; IEA.
 DE GO:0004295; F:trypsin activity; IEA.
 DE GO:0006508; P:proteolysis and peptidolysis; IEA.
 DE InterPro; IPR009003; Cys_Ser_trypsin.
 DE InterPro; IPR000742; EGF_2.
 DE InterPro; IPR006209; EGF_like.
 DE InterPro; IPR000083; Fibrinctnl.
 DE InterPro; IPR000562; FN_Type_II.
 DE InterPro; IPR006210; IEGF.
 DE InterPro; IPR000001; Kringle.
 DE InterPro; IPR001254; Peptidase S1.
 DE InterPro; IPR001314; Peptidase_S1A.
 DE Pfam; PF00008; EGF; 2.
 DE Pfam; PF00039; fn1; 1.
 DE Pfam; PF00040; fn2; 1.
 DE Pfam; PF00051; kringle; 1.
 DE Pfam; PF00089; trypsin; 1.
 DE PRINTS; PR00722; CHYMOTRYPSIN.
 DE PRINTS; PR00013; ENTYPSEII.
 DE PRINTS; PR00018; KRINGLE.
 DE ProDom; PD000995; FN_Type_II; 1.
 DE ProDom; PD000395; Kringle; 1.
 DE SMART; SM00181; EGF; 2.
 DE SMART; SM00059; FN2; 1.

SMART; SM00130; KR; 1.
 SMART; SM00020; Tryp_Spc; 1.
 PROSITE; PS00022; EGF_1; 2.
 PROSITE; PS01186; EGF_2; 1.
 PROSITE; PS01253; FIBRONECTIN 1; 1.
 PROSITE; PS00023; FIBRONECTIN 2; 1.
 PROSITE; PS00021; KRINGLE 1; 1.
 PROSITE; PS00070; KRINGLE 2; 1.
 PROSITE; PS02040; TRYPSIN_DOM; 1.
 PROSITE; PS00134; TRYPSIN_HIS; 1.
 PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
 KW Kringle; Protease; Serine protease.
 SQ SEQUENCE 653 AA; 70553 MW; F818D90174ED6DD CRC64;
 Query Match 31.7%; Score 716; DB 11; Length 653;
 Best Local Similarity 36.3%; Pred. No. 6.6e-59;
 Matches 154; Conservative 58; Mismatches 162; Indels 50; Gaps 8;
 QY 5 HQVPSNCDLNGGTCVSNKYFSNIHW-----CNCPKKFGGQHCIDKSKTCYENGH 56
 DB 239 HTACLSSPLNGGTC-----HLIVGTGTSVCTCPLGVAGRFNCIVPTEHCFLGNGT 289
 QY 57 FYRGKASDTWGRCLPWN SATVLQOTVHAHRS DALQLGLGKHNYCNPDRRRPWCYVQ 116
 DB 290 EYRGVASTAASGLSCLAWNSDLLYQELHVDSEKAVOLGGLGFPSCRNPDDEDEKFWCT 349
 QY 117 VGLKPLVQECMVHDC-----ADGKLKFCQCK-----TLRPFKII 152
 DB 350 KDNALSWEYCYRLTACESLARVHSQSPEILAAIPESAPAVRPTCGKRHKRTFLRPR--II 407
 QY 153 GGEFTIENQWPFAAYRHRHGGSVTVYCGSLSPCWVISAHCFTDYPKEDYIVYLG 212
 DB 408 GSSSLPGSHPTAAIV---IGNS---FCAGSLVHTCWVWSAAHCFANSPDRSITVVLG 461
 QY 213 RSLNSNTQCEMKFEVENLILHKDYSADTLAHNDIALKIRSKEGRCAPQRTIQTICL 272
 DB 462 QHFFNRRTDVTQTGIEKYVPTLYSVFNPNNH-DLVILRLKKKGERCAVESQFVQPICL 520
 QY 273 PSYNDPQFTSCITGSGKENS DYLPEOLKMTVVKLISHRECQPHYVGSVTTKML 332
 DB 521 PEAGSFPTGHKKQIAGHGHNDENVSSYNSLLBALPVLVADHKCSPEVYGAISPNNML 580
 QY 333 CAADPQWKTSCOGDGGPLVCSLQGRMTLTGIVSWGRGCALDKPGVYTRVSHFLPWI 392
 DB 581 CAGYFDSKSDACOGDGGPLVCKNGVAYLYGIISWGDGCRVKNKPGVYTRVSNVNI 640
 QY 393 SHTK 396
 DB 641 DRIR 644

RESULT 13
 O97507 PRELIMINARY; PRT; 616 AA.
 AC O97507;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE FXII.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RE SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Takahashi T.; Kihara T.;
 RT "porcine liver factor XII".
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DB EMBL; AB022426; BAA37148.1; -.

183 CLHGRCLUE---VEGRLCHCPVGTGTFCDVDTKASCDGRGLSYRGLARTTILSGAPCQ 239
73 PWSATVLQOQTY-HAHRSDALQGLGKHNCRNPDNRERRPWCYVOVGLKPLVQECMVHDC 131
240 PWAS-----EATYRNVTAEQARNWGLGGHAFCRNPDNDIRPWCFFVLRDRLSWYCDLAQC 295
132 -----ADGKL----- 136
296 QTPQAAPPTVSPRLHVPLMPAAPPKPOPTTRTPPQSQTPGALPAKREQPPSLTRNG 355
137 KFGCGQ---KTLRPRFKIIGEFTTIENQPFALYRHRGGSVTVYVCGSLISPCWVIS 193
356 PLSCGQRURKLSWTRVVGGLVALRGANHPIALYWGHS-----FCAGSLIAPCWLT 409
194 ATHCFIDYPKKEDYVYVLRGRSLNNTQGMKFEVENLILHKDYADTLAHHNDIALK 253
410 AAHLQDRPAPEDLTVLVQGRNRHSCBPCQTLAVRSYRLHEAFS--PVSQYQHDALLRL 467
254 R-SKEGRCAQPSRTIOTICLPSMYNDPQFGSCITGFKENSVDYLYPEQLXMTVVKLI 312
468 QEDADGSCALLSPYQVPLPSGGAARSETTLCQVAGCGHQFEGAEYASFLOEAQVFFL 527
313 SHRECQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQG---RMTLTGIVSWG 369
528 SLERCSADVHGSSILPGMLCAGFLEGGTDACQDGGPLVCEQQAARLTLQGIISWG 587
370 RGCALKDKGYNTRYSHPLWIRSH 395
588 SGCGDRNKPVGYYTDVAYYLAWIREHT 613

RE LT 15
O3 27
IE O35727 PRELIMINARY; PRT; 597 AA.
AC O35727;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Factor XII.
GN F12
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX [1]
RE SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Schloesser M., Schwager S., Engel W.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DE EMBL; X99571; CAA67891.1; -;
DE HSSP; P00760; 1A07.
DE MEROPS; S01.211; -;
DE MGD; MGI:1891012; F12.
DE GO; GO:0005576; C:extracellular; IEA.
DE GO; GO:0004263; F:chymotrypsin activity; IEA.
DE GO; GO:0008233; F:peptidase activity; IEA.
DE GO; GO:0004295; F:trypsin activity; IEA.
DE GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DE InterPro; IPR009003; Cys Ser trypsin.
DE InterPro; IPR006209; EGF-like.
DE InterPro; IPR000083; Fibinctnl.
DE InterPro; IPR000562; FN_Type_II.
DE InterPro; IPR006210; IEGF.
DE InterPro; IPR000001; Kringle.
DE InterPro; IPR001254; Peptidase S1.
DE InterPro; IPR001314; Peptidase_S1A.
DE Pfam; PF00008; EGF; 2.
DE Pfam; PF00039; fn1; 1.
DE Pfam; PF00040; fn2; 1.
DE Pfam; PF00051; kringle; 1.
DE Pfam; PF00089; trypsin; 1.

DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0013; ENTPEI1.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;
Query Match 29.6%; Score 668; DB 11; Length 597;
Best Local Similarity 35.6%; Pred. No. 2.1e-54;
Matches 152; Conservative 61; Mismatches 154; Indels 60; Gaps 13;
QY 13 CLNGGTCVSNKYFSNIHWCNCPKKEGGQHCIDKSKTCYEGNGHYRGKASTDTMGRPCL 72
DB 183 CLNGGSC---LVEDHPLCRCPFTGYTGYFCOLDLWATCYEGRLSYRQAGTTQSGAPCQ 239
QY 73 PWSATVLQOQTY-HAHRSDALQGLGKHNCRNPDNRERRPWCYVOV----- 118
DB 240 RM-----TVEATYRNTEKQALSWGLGHAFCEPNPDNTEPWCFFVMSGDELSDYDGLBQC 295
QY 119 -----LKPLV-----QE-----CMVH-----DCADGKLKFO-----CGQ---KTLRPRFKI 151
DB 296 QTPTFAPLVPSQESPSQAPSLSHAPNDSTDHQTSLSKTNTMCGQGRFRKGLSPMRV 355
QY 152 IGGEFTTIENQPFALYRHRGGSVTVYVCGSLISPCWISATHCIFDYPKEDYIVVL 211
DB 356 VGLVALPQSHFYIALYGNW-----FCAGSLIAPCWLTAAHCLNRPAPELTIVL 409
QY 212 GSRSLNNTQGMKFEVENLILHKDYADTLAHHNDIALKIR-SKEGRCAQPSRTIQT 270
DB 410 GQDRHNQSCWCQTLAVRSYRLHEGFSITYGH--DLALLRQESKTNCAILSPHVQPV 467
QY 271 CLPSMYNDPQFGSCITGFKENSTDYLYPEQLXMTVVKLISHRECQCPHYGSEVTK 330
DB 468 CLPSGAAPSETVLCVAGWHQLEGAEYSTFLOEAQVFFIALDRCSNINVHGDAILPG 527
QY 331 MLCAADPQWKTDSCQDGGPLVC---SLQGRMTLTGIVSWGRGALCKPKGVYTRVSHF 387
DB 528 MLCAGFLEGGTDACQDGGPLVCEGTAETHQLTLRGVISWGGCGDRNKPVGYYTDVANY 587
QY 388 LPKIRSH 394
DB 588 LAWIQKH 594
RESULT 16
Q80YC5 PRELIMINARY; PRT; 609 AA.
ID Q80YC5
AC Q80YC5
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Similar to coagulation factor XII (Hageman factor) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]

SEQUENCE FROM N.A.
STRAIN=FVB/N; TISSUE=Liver;
Strausberg R.; to the EMBL/GenBank/DBJ databases.
Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
EMBL; BC049867; AAH49867.1; -
GO; GO:0005576; C:extracellular; IEA.
GO; GO:0005509; F:calcium ion binding; IEA.
GO; GO:0004263; F:chymotrypsin activity; IEA.
GO; GO:0004295; F:trypsin activity; IEA.
GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
InterPro; IPR009003; Cys_Ser_trypsin.
InterPro; IPR01881; EGF_Ca.
InterPro; IPR006209; EGF-like.
InterPro; IPR000083; Fibrinctn1.
InterPro; IPR000562; FN_Type_II.
InterPro; IPR006210; IEGF.
InterPro; IPR000001; Kringle.
InterPro; IPR001254; Peptidase_S1.
InterPro; IPR001314; Peptidase_S1A.
Pfam; PF00008; EGF_2.
Pfam; PF00039; fn1; 1.
Pfam; PF00040; fn2; 1.
Pfam; PF00051; kringle1; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00013; FNTYPEII.
PRINTS; PR00018; KRINGLE.
ProDom; PD000995; FN_Type_II; 1.
ProDom; PD000395; Kringle; 1.
SMART; SM00181; EGF_2.
SMART; SM00179; EGF_Ca; 2.
SMART; SM00058; FN1; 1.
SMART; SM00059; FN2; 1.
SMART; SM00130; KR; 1.
SMART; SM00020; Tryp_Spc; 1.
PROSITE; PS00022; EGF_1; 2.
PROSITE; PS01185; EGF_2; 1.
PROSITE; PS01253; FIBRONECTIN_1; 1.
PROSITE; PS00023; FIBRONECTIN_2; 1.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS00070; KRINGLE_2; 1.
PROSITE; PS00240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
NON_TER; 1
SEQUENCE 609 AA; 6783 MW; DF97D4D2369B6D2 CRC64;
Query Match 29.6%; Score 668; DB 11; Length 609;
Best Local Similarity 35.6%; Pred. No. 2.le-54;
Matches 152; Conservative 61; Mismatches 154; Indels 60; Gaps 13;
13 CLNGGTCVSKYFSNTHWCNCPKFGQHCIDKSKTCYEGNGHFRGKASTDTMGRPCL 72
195 CLNGGSCLE---LVEDHPLCRCPGTGYTGFDLDLWATCYEGRGLSYRGQAGTTQSGAPCQ 251
73 PWSATVLOQTY-HAHRSDALQLGLGKHNCRPNRRRRCVYVQV-----118
252 RW-----TVEATYRWTEKQALSWGLGHAFCRPNPDNDTRFCVWSGDRLSWDYCYGLEQC 307
119 -----LKPVLV-----QEV-----CMVH-----DCADGKLFQ-----CGQ-----KTLRPRFKI 151
308 QTTFAPLVVPESSQSPAPSLSEAPNDSTDHQTSKNTMTGCGQRFKGLSSFMFV 367
152 IGGFTTIEQNPFAATYRHRGGSVTVYCGGSLISPCWVISATHCIDYPKKEDIVVL 211
368 VGLUVALPGSHPIYALYWN-----FCAGSLIAPCWVIAHCLQRPAPELTIVL 421
212 GRSLNNTQGEKMFVENILHKVYSADTLAHNDIALKIR-SKEGRCAQPSRTIQT 270
422 GQDRHNSCEWCQTLAVRSYRLHEGFSSITYQH--DLALLRLQESKNTSCAILSPHVQV 479
271 CLPSMYNDPQGTSCETIGTKENSTIDYLPOLKMTVVKLISHRSCQPHYVGSVYTK 330
Db 480 CLPSGAAPPSETVLCVAGHGQHEGAEYSTFLQEAQVFIALDRCSNVHGDAILPG 539
Qy 331 MLCADPQWKTDSQGDGGPLVC---SLQGRMTLTGIVSWGRGCAKLDKPGVYVYSHF 387
Db 540 MLCAGFLGGTDAQCGDGGPLVCCEGTAEHQLTLRGVSWGSGCGDRNKPQVYTDVANY 599
Qy 388 LPWIRSH 394
Db 600 LAWIQKH 606
RESULT 17
Q8K0D2 PRELIMINARY; PRT; 517 AA.
ID Q8K0D2
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC031775; AAH31775.1; -
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF_2.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;
Query Match 29.4%; Score 664.5; DB 11; Length 517;
Best Local Similarity 36.9%; Pred. No. 3.7e-54;
Matches 154; Conservative 57; Mismatches 163; Indels 43; Gaps 14;
Qy 13 CLNGGTCVSKYFSNTHWCNCPKFGQHCIDKSKTCYEGNGHFRGKASTDTMGRPCL 72
Db 116 CQNGVCSRRHRRSRF-TCACPDQYKGFCEIGPD-DCYVGDGYSYRGKYSKTVNQNPCL 173
Qy 73 PWSATVLOQTYHAHRSDALQLGLGKHNCRPNRRRRCVYVQVGLKPL-----VQEC 126

DB 174 YWNSHLLQTYNMFEDASTHGAENFCNPDGDKWCFVKVNSEKVKWEYCDVTV 233
QY 127 MVDCADCKLKF-----CCQKTLRPP--KIIGBETTIENQWFAAIY---- 169
DE 234 FVPTNPVBSLLEPVMELPGFSCGTEVAEHAHVRIYGGFKTAGKHPQVLSQTSUP 293
QY 170 ---RRHRGGSVTVYVCGSLISPCWVISAHCFFIDYPKKEDYIYVLSRSLNSNTQGMKP 226
DB 294 LTISMPQG---HFCGGLIHPQWLVTAHC-TDINTKHLKV-LGDQDLKKTESHEQTF 347
QY 227 EVENLIHKYSDATLAHNDIALLKRSKRGCAQPSRTIOTICLPSMYNDP-QFGTSC 285
DE 348 RVEKILKYQSNRDEIPHNDIALLKLPVGGHCALESRYVKTCLPS---DFFPSGTSC 404
QY 286 EITGFGKENSVDYLPQKXMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWK-TDSC 344
DE 405 HISGNGVTEGE--GSQLLDAKVKLIANPLNSRQLYDHTIDDSMICAGNLQKPGSDTC 462
QY 345 QGDSGGPLVCSLQGRMTLTGIVSGRCALKDPGVTVTRVSHFLPWIRSHTKKEG 401
DB 463 QGDSGGPLTCEKDTGVYVGVISWGCEG--KRPGVYTVQTKFLNWKITMTHREAGL 517
RE 1T 18
Q1 20
ID Q14520 PRELIMINARY; PRT; 560 AA.
AC Q14520; O00663;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DI 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE HGF activator like protein (Hyaluronan binding protein 2).
OE Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
SEQUENCE FROM N.A.
RA Kitamura N.
RP Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RL
RN
RF
RX MEDLINE=96425001; PubMed=9827452;
RA Chai-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
RA Tomita M.;
RT "Purification and characterization of a novel hyaluronan-binding
RT protein (PHBP) from human plasma: it has three EGF, a kringle and a
RT serine protease domain, similar to hepatocyte growth factor
RT activator.",
RL J. Biochem. 119:1157-1165(1996).
RN [3]
SEQUENCE FROM N.A.
RC TISSUE=Colon, and Kidney;
RA Strausberg R.;
RI Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DE EMBL; D49742; BAA08576.1; -.
DE EMBL; S83182; AAB46909.1; -.
DE EMBL; BC031412; AAH31412.1; -.
DE PIR; JC4795; JC4795.
DE HSSP; P00763; LDPO.
DE MEROPS; S01.033; -.
DE Genew; HGNC:4798; HARP2.
DE GO; GO:0005615; C:extracellular space; TAS.
DE GO; GO:0005539; F:glycosaminoglycan binding; TAS.
DE GO; GO:0007155; P:cell adhesion; TAS.
DE InterPro; IPR009003; Cys Ser trypsin.
DE InterPro; IPR006209; EGF like.
DE InterPro; IPR000001; Kringle.
DE InterPro; IPR001234; Peptidase_S1.
DE InterPro; IPR001314; Peptidase_S1A.
PFam; PF00008; EGF; 3.

DR PFam; PF00051; kringle; 1.
DR PFam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;
Query Match 28.9%; Score 651.5; DB 4; Length 560;
Best Local Similarity 36.4%; Pred. No. 6.9e-53;
Matches 152; Conservative 58; Mismatches 161; Indels 47; Gaps 14;
QY 13 CLNGSTCVSNKYFNHWCNCPKKGGOHCEIDSKTCYEGNGHFYRGKASTDTWGRPCL 72
DB 159 CQNGATCSRHRKRSKF-TCACPDQFKGKCEIG-SDDCYVGGYSGYRGKMRVTNQHACL 216
QY 73 PWSATVLTQTYHAHRSDALQLGLGKHNCRPNRRPWCYVQVGLKPLVQE-CMVHDC 131
DB 217 YWNSHLLQENYMFEDASTHGAENFCNPDGDKWCFVKVNSEKVKWEYCDVSA 276
QY 132 ADGKLKF-----CCQKTLRPP--FKIIGBETTIENQWFAAIYRHR 173
DB 277 SAQDVAYPEESTPSTKLPGFDSGKTEIAERIKRIYGGFKTAGKHPQVSAIQ---- 332
QY 174 GGSVT-----YVCGSLISPCWVISAHCFFIDYPKKEDYIYVLSRSLNSNTQGMK 225
DB 333 -SSLPTTISMPQGHFCGGLIHPQWLVTAHC-TDI-NTRHLKVVLGDQDLKKEEFHQS 389
QY 226 FEVENLIHKYSDATLAHNDIALLKRSKRGCAQPSRTIOTICLPSMYNDPQF--GT 283
DB 390 FRVEKIFKYSYNERDEIPHNDIALLKLPVGGHCALESRYVKTCLP----DGSFPPSGS 445
QY 284 SCEITGFGKENSVDYLPQKXMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWK-TD 342
DB 446 ECHISGNGVTEGE--GSQLLDAKVKLIANPLNSRQLYDHTIDDSMICAGNLQKPGD 503
QY 343 SCQDSGGPLVCSLQGRMTLTGIVSGRCALKDPGVTVTRVSHFLPWIRSHTKKEG 400
DB 504 TCQDSGGPLTCEKDTGVYVGVISWGCEG--KRPGVYTVQTKFLNWKITKSES 559
RESULT 19
Q1TVAS PRELIMINARY; PRT; 157 AA.
AC Q1TVAS;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21071388; PubMed=11204721;
RA Balcerzak D., Quereguesser L., Dixon W.T., Baracos V.E.;
RT "Coordinate expression of matrix-degrading proteinases and their
RT activators and inhibitors in bovine skeletal muscle.",
RL J. Anim. Sci. 79:94-107(2001).
CC -|- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DE EMBL; AF144761; AAD30301.1; -.
DR HSSP; P00749; IURK.

GO: GO:0016301; F:kinase activity; IEA.
InterPro; IPR006209; EGF-like.
InterPro; IPR000001; Kringle.
Pfam; PF00051; kringle; 1.
PRINTS; PR00018; KRINGLE.
ProDom; PD000395; Kringle; 1.
SMART; SM00130; KR; 1.
PROSITE; PS00022; EGF_1; 1.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS00070; KRINGLE_2; 1.
Glycoprotein; Kinase; Kringle.
NON_TER 157 157
SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBB7 CRC64;

Query Match 28.3%; Score 638; DB 6; Length 157;
Best Local Similarity 70.7%; Pred. No. 2.7e-52;
atches 111; Conservative 14; Mismatches 24; Indels 8; Gaps 1;

13 CLANGTCVSNKYNSNIHWCNCPKFGGQHCEIDKSKTCYBNGHGFYRGKASTDTMGRPCL 72
1 CLANGKCVTVKYSNIORCSGPKFGQEHCEIDTSKTCYGNHSGYRGKANDLSRPL 60

73 PWSATVLOQTYAHRSALQLGLGKHNYCRNPNRRRPMCYVQVGLKPLVQECMYHDC 132
61 AWPSTVLLKMYAHRSDAIQGLGKHNYCRNPNQRRPWCYVQVIGLKQFVQCMVQDCS 120

133 DGKL-----KFQCGKTLRPREKLIIGFEFTIEN 161
121 VGKSPSPSREKRFQCGKALRFRKIVGQVTNAEN 157

ULT 20
587 PRELIMINARY; PRT; 128 AA.
O97587;
A O97587;
01-MAY-1999 (TrEMBLrel. 10, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Urokinase (Fragment).
Oryctolagus cuniculus (Rabbit).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NCBI_TaxID=9986;
[1]
SEQUENCE FROM N.A.
STRAIN=New Zealand White;
MEDLINE=95057575; PubMed=9837780;
Reno C.; Boykiw R.; Martinez M.L.; Hart D.A.;
"Temporal alterations in mRNA levels for proteinases and inhibitors
and their potential regulators in the healing medial collateral
ligament.";
Biochem. Biophys. Res. Commun. 252:757-763(1998).
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
EMBL; AF069711; AAC95003.1; -.
HSP; P00749; IEJN.
MEROPS; S01.231; -.
GO; GO:0004263; F:chymotrypsin activity; IEA.
GO; GO:0016301; F:kinase activity; IEA.
GO; GO:0008233; F:peptidase activity; IEA.
GO; GO:0004295; F:trypsin activity; IEA.
GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
InterPro; IPR009003; Cys.Ser trypsin.
InterPro; IPR001254; Peptidase_S1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
SMART; SM00020; TRYP_SPC; 1.
PROSITE; PS00240; TRYPIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
Hydrolase; Kinase; Protease; Serine protease.
NON_TER 1 128

SQ SEQUENCE 128 AA; 14328 MW; 1BC7ED30E071A06D CRC64;

Query Match 25.9%; Score 585; DB 6; Length 128;
Best Local Similarity 84.4%; Pred. No. 2.1e-47;
Matches 108; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 164 WFAALYRRHSGSVTVCGSLISPCWVISAHGFIDYPKKEDYIVLGRSLNSNTQGE 223
Db 1 WFAALYRRHSGSVTVCGSLISPCWVISAHGFIDYPKKEDYIVLGRSLNSMTFGE 60

QY 224 MKFEVENILHKDYSADTLAHNDIALIKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGT 283
Db 61 MKFEVEQLILHEGYRADTLAHNDIALIKILSNNGCAQPSRSIQTICLPPWADPNFGT 120

QY 284 SCEITGFG 291
Db 121 SCEITGFG 128

RESULT 21
Q9ROW3 PRELIMINARY; PRT; 812 AA.
AC Q9ROW3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
GN PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bangert K.; Johnsen A.H.; Thorsen S.;
"Rat plasminogen: cDNA and gene structure";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91250378; PubMed=1645711;
RA Kanalas J.J.; Makker S.P.;
"Identification of the rat Heymann nephritis autoantigen (GP330) as a
receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829(1991).
CC -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; AJ242649; CAB46014.1; -.
DR HSP; P00747; 1PMK.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004283; F:plasmin activity; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys.Ser trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_SIA.
DR InterPro; IPR003966; Peptidase_SIA_pr.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.

```

CO; GO:0008233; F:peptidase activity; IEA.
GO; GO:0003809; F:thrombin activity; IEA.
GO; GO:0004295; F:trypsin activity; IEA.
GO; GO:0007596; F:blood coagulation; IEA.
GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR DR InterPro; IPR009003; Cys Ser trypsin.
DR DR InterPro; IPR000001; Kringl_
DR DR InterPro; IPR003014; PAN.
DR DR InterPro; IPR003509; Pan_app.
DR DR InterPro; IPR001254; Peptidase_S1.
DR DR InterPro; IPR001314; Peptidase_S1A.
DR DR InterPro; IPR003966; Peptidase_S1A_pr.
DR DR Pfam; PF00051; kringle; 5.
DR DR Pfam; PF00024; PAN; 1.
DR DR Pfam; PF00089; trypsin; 1.
DR DR PRINTS; PR00722; CHYMOTRYPSIN.
DR DR PRINTS; PR00018; KRINGLE.
DR DR PRINTS; PR00105; PROTHROMBIN.
DR DR ProDom; PD000395; Kringle; 5.
DR DR SMART; SM00130; KR; 5.
DR DR SMART; SM00473; PAN_AP; 1.
DR DR SMART; SM00020; TRY SP; 1.
DR DR PROSITE; PS00021; KRINGLE_1; 5.
DR DR PROSITE; PS00070; KRINGLE_2; 5.
DR DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR DR Glycoprotein; HydroLase; Kringle; Protease; Serine protease; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 810 PLASMINOGEN.
SQ SEQUENCE 810 AA; 90555 MW; B05C7D4B0D020B3C CRC64;

Query Match 22.1%; Score 498; DB 4; Length 810;
Best Local Similarity 33.4%; Pred.No.3.7e-38;
Matches 136; Conservative 45; Mismatches 170; Indels 56; Gaps 15;

QY 10 NCDCLNGTGVSNKYFSNIHWCNPKFGGQ-----HCEIDKSKTCYEGNGH 56
Db 428 NPADKGPWCFTDPSVRWEYCNLKCKSGTEASVVAPPVLLPDVETPSEEDCMFGNGK 487

QY 57 FYRGKASTDTWGRPCLPNNSATVLQOTYHAHR---SDALQGLGKHNYCRNP-RRRRP 111
Db 488 GYRGKRAITVTGTPQODWAA-----QEPHRHSITFPENPRAGLEK-NYCRNPDGDVGGP 541

QY 112 WCYOVGGLKPLVQECMHWDCAAGLKFCQGKTLRPR---FKIIGETTTIENQWFAAI 168
Db 542 WCYT-TNPRKLYDYCDVPQCA--APSPFCGKQPVEPKPCRGVVGCGVAHPHSPWQVSL 598

QY 169 YRRHGGSVTVYCGGSLISPCWVTSATCFIDYKKEDYIVVLGRSLRNSNTQGMKFEV 228
Db 599 --RTAFG--MFPCGGTLLSPWLVTAACHLEKSPSPSSKYKVLGAHQEVNLEPHVQOEIEV 654

QY 229 ENLIHKDYSADTLAHHNDIALKIRSKRGCAQPSRTIQITCLPSMYNDPQF----GTS 284
Db 655 SELFLEPT-----RKDIALKLSSP-----AVITDKVI PACLPS----PNYVVADRTE 698

QY 285 CEITGFGKENSTDYLYPEQLKMTVVKLLISHRECOQPHYVGSEVTTKMLCAADPQWKTDSC 344
Db 699 CFVTGWGTQGT--FGAGLLKEAQLPVLENKVCNRYEFLNGRVQSTELCAGHLAGGTDSC 756

QY 345 QDGSGGPLVCSLQGRMTVTGVISWGRGCALKDKKPGVYTRVSHFLPMI 391
Db 757 QDGSGGPLVCFEKDKYILQGVTSWGLGCARPKNKPGVYVRVRSREFTWI 803

RESULT 23
Q95M89 PRELIMINARY; PRT; 103 AA.
ID Q95M89
AC Q95M89
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen activator urokinase (Fragment).

```

PLAU.
C Equus caballus (Horse).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
C NCBI_TaxID=9796;
C [1]
F F MEDLINE=21314992; PubMed=11421942;
F Shubtowski D.M., Venta P.J., Douglass C.L., Zhou R.-X., Ewart S.L.;
F "Polymorphism identification within 50 equine gene-specific sequence
F tagged sites.";
F Anim. Genet. 32:78-78(2001).
C -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
C EMBL; AY008806; AA14840.1;
D EMBL; AY008803; AA14840.1; JOINED.
I GO; GO:0004263; F:chymotrypsin activity; IEA.
I GO; GO:0016301; F:kinase activity; IEA.
I GO; GO:0008233; F:peptidase activity; IEA.
I GO; GO:0004295; F:trypsin activity; IEA.
I GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
I InterPro; IPR009003; Cys Ser trypsin.
I InterPro; IPR001254; Peptidase_S1.
I Pfam; PF00089; trypsin; 1.
I PRINTS; PR00722; CHYMOTRYPSIN.
I SMART; SMO0202; Tryp_Spc; 1.
I PROSITE; PS0240; TRYPSIN_DOM; 1.
I PROSITE; PS00134; TRYPSIN_HIS; 1.
I Hydroxylase; Kinase; Protease; Serine protease.
I NON_TER 1 103
F SEQUENCE 103 AA; 11525 MW; 05739514F6331180 CRC64;
E
S
Query Match 21.9%; Score 495; DB 6; Length 103;
Best Local Similarity 87.4%; Pred. No. 5.4e-39;
Matches 90; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
164 WFAIYRRHGGSVYVCGSLSPCWVISAATHCFIDYPKEDYIVYVGLSRSLNQTQGE 223
1 WFAIYRRHGGSVYVCGSLSPCWVISAATHCFIDYPKEDYIVYVGLSRSLNQTQGE 60
224 MKEFEVENLIHKYSDTLAHNDIALKIRSKEGRCQAQPSRT 266
61 MKEFEVENLIHKYSDTLAHNDIALKIRSKEGRCQAQPSRS 103
ULT 24
EF3
Q96EF3 PRELIMINARY; PRT; 300 AA.
Q96EF3
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Hypothetical protein (Coagulation factor XII) (Hageman factor).
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Lung;
Strausberg R.;
Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
Kalinine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
Phelan M., Farmer A.;
Cloning of human full-length cDNAs in BD Creator(TM) System Donor
vector.";
Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
EMBL; BC012390; AA12390.1; -.

DR EMBL; BT007350; AAP36014.1; -.
DR HSSP; P00761; IAN1.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SMO0202; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; Hydrolase; Protease; Serine protease.
SQ SEQUENCE 300 AA; 32216 MW; FFC2BDF9382F638A CRC64;
Query Match 21.9%; Score 493.5; DB 4; Length 300;
Best Local Similarity 37.3%; Pred. No. 2.8e-38;
Matches 112; Conservative 46; Mismatches 119; Indels 23; Gaps 7;
QY 103 RNPENRRPWCYVQVGLKPLVQECWHDGADGKLKFCGQ---KTLAPRPFKIIIGERTTI 159
DB 15 RTPQSQTP-----GALPAKREQPPSLTRNGPL--SCGQLRKSLSMTVRVGLVALR 66
QY 160 ENQPFWFAIYRRHGGSVYVCGSLSPCWVISAATHCFIDYPKEDYIVYVGLSRSLN 219
DB 67 GAHFVIAALYNGHS-----FCAGSLIAPCWVLTAAHCLQDRPAPEDLTIVVGLQERRNHS 120
QY 220 TQGEKKEVENLIHKYSDTLAHNDIALKIR-SKEGRCQAQPSRTIQTICLPSMYND 278
DB 121 CEPQTLAVRSYRLHEAFS--PVSYQHDLLALRIQEDADGSCALLSPYVQVCLPSGAAR 178
QY 279 PQFTSCBITGKGNSTDYLPQLKMTVVKLSHRECCQPHYVGGSEVITKMLCAADPQ 338
DB 179 PSEITLCQVAGHQFEGAEYASFLQEAQVFPFSLERCSAPDVHGSILPGMLCAGFLE 238
QY 339 WKTSCQDSDGGLVCSLQG---RMTLTGVSWGRGKALDKDKPGVYTRVSHFLPWIRSH 395
DB 239 GGTDACQDSDGGLVCSLQG---RMTLTGVSWGRGKALDKDKPGVYTRVSHFLPWIRSH 298
RESULT 25
Q46506 PRELIMINARY; PRT; 454 AA.
AC Q46506;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Apolipoprotein a (Fragment).
GN B4BAP0A.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RA Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation is Associated with Deletion of a Single Exon in a Null
RT Allele."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF029691; AAB97886.1; -.
DR HSSP; P00747; 2PK4.
DR MEROPS; S01.999; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; P:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro: IPR009003; Cys Ser trypsin.
 DR InterPro: IPR000001; Kringle_1.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00051; Kringle; 2.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00118; KRINGLE.
 DR ProDom: PD000395; Kringle; 2.
 DR SMART: SM00130; KR; 2.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS00021; KRINGLE_1; 2.
 DR PROSITE: PS00070; KRINGLE_2; 2.
 DR PROSITE: PS0240; TRYPSIN_DOM; 1.
 DR PROSITE: PS0135; TRYPSIN_SER; 1.
 DR Glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;
 KW Serine protease.
 FT NON_TER 1
 SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;
 37y Match 21.8%; Score 493; DB 6; Length 454;
 st Local Similarity 33.8%; Pred. No. 5.3e-38;
 tches 127; Conservative 51; Mismatches 134; Indels 64; Gaps 14;
 Qy 50 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQ--QTYHAHRS DALQLGLGKHNVCNPN 107
 Df 102 CYHGDGSGYRGSTFTVTGRTCSWSSMTPHQKRTPEHPDGLTV-----NICRNPDA 156
 Qy 108 RRRPWCYVQVGLKPLV--QECMVHDCAD-----GKLPQCGQ 442
 Db 157 DTGFWCFI---MDPSVAREYCNLFRCSDTGTGVTPLTVIPSLARSQASSPDCGK 213
 Qy 143 KTLAPR---FKIIGETTTNQWFAIYRHRGGSVTVCGSLISPCWVLSATHCFI 199
 Db 214 PQVEFKCPGVRGVGCAHAHSPWQVSL--RTRFGK--HFCGGLTSLPWWVLTARCLE 269
 Qy 200 DYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKGR 259
 Db 270 MSPSPSSKVLG-----AHQEVNLESHVQIEVSLFSPTGA---DIALKL-----SR 317
 Qy 260 CAQSPRTIQTCLSMYNDPOF-----GTSCEITGFGKENSTDIYLPQLKMTVVKLISHR 315
 Db 318 PAITDKVIPACLPFS---PNYITAWTECYITGWGETGT--FGAGLLREARLPVIENT 371
 Qy 316 ECQPHYVYGVETTKMLCAADPQWKTDSCQDSCGGLVCSLQGRMTLTGIVSWGRCALK 375
 Df 372 VCNRYEPLNGRVKSTELCAGHLAGTDSQDSCGSGPVVCFDKKYLIRGITSNGPGCARP 431
 Qy 376 DKPGYVTVSHFLPMI 391
 Df 432 NKPGYVTVSVSFVTWI 447

RE Q8 26
 ID Q8ABV0 PRELIMINARY; PRT; 429 AA.
 AC Q8ABV0;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Plasminogen precursor (Fragment).
 DE Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 RN [1]
 RE SEQUENCE FROM N.A.
 RA Hanumanthaiah R., Day K., Jagadeeswaran P.;
 RI "Comprehensive analysis of blood coagulation pathways in teleostei:
 RI Evolution of coagulation factor genes and identification of zebrafish
 RI factor VIII".
 RI Blood Cells Mol. Dis. 0:0-0(2002).

DR EMBL: AF515276; AAN71006.1;
 DR GO: GO:0004263; F:chymotrypsin activity; IEA.
 DR GO: GO:0004295; F:trypsin activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00051; Kringle; 2.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00118; KRINGLE.
 DR ProDom: PD000395; Kringle; 2.
 DR SMART: SM00130; KR; 2.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS00070; KRINGLE_2; 2.
 DR PROSITE: PS0240; TRYPSIN_DOM; 1.
 DR PROSITE: PS0134; TRYPSIN_HIS; 1.
 DR PROSITE: PS0135; TRYPSIN_SER; 1.
 FT NON_TER 1
 SQ SEQUENCE 429 AA; 47556 MW; 9A580A214A549C12 CRC64;
 Query Match 21.8%; Score 491.5; DB 13; Length 429;
 Best Local Similarity 35.3%; Pred. No. 6.8e-38;
 Matches 127; Conservative 37; Mismatches 155; Indels 41; Gaps 13;
 Qy 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQ---TYHAHRS DALQLGLGKHNVCN 104
 Db 100 KDCXNGNGAEYRGSTFTVTGVTQAMRSMTPHQAFTPTHDPKGL-----SNQCRN 154
 Qy 105 PDNR--RRPWCYVQVGLKPLVQECMVHDCADKGLKFCQCKTLRPR---FKIIGETTTIE 160
 Db 155 PSDVNGPWCYVTDPSKKW--DYCQIPDCE---SLKCGQATPKRCFCGRIVGCVSKPH 209
 Qy 161 NOPWFAIYRHRGGSVTVCGSLISPCWVLSATHCFIDYPKKEDIYVIGSRSLNSNT 220
 Db 210 SWPW--QISLTRG--KIHFCGGLTIDPQWVTAACHLERSDPSAVKIMLGITERATE 265
 Qy 221 QCEMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCAQSPRTIQTICLPSMYNDPQ 280
 Db 266 SSKQERDVKLI---KGPAQT-----DIALKL-----DRPALINDKVPCLPEKDYIVP 313
 Qy 281 FQTSCEITGFGKENST---DYLPEQLKMTVVKLISHRECOOPHYGVSEVTTKMLCAADP 337
 Db 314 SNTCEYVTVGWGTQDTQGGY-----LKEITGFVIENKVCNRPFLNVRKVDHENCAGNI 368
 Qy 338 QMKTDSCQDSCGGLVCSLQGRMTLTGIVSWGRCALKDKPGYVTVSHFLPWRSHTKTE 397
 Db 369 EGNDSQDSCGGLVCSLQGRMTLTGIVSWGRCALKDKPGYVTVSHFLPWRSHTKTE 428

RESULT 27
 O46507 PRELIMINARY; PRT; 334 AA.
 ID O46507
 AC O46507;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Plasminogen (Fragment).
 GN BABEPSPG.
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopithecinae; Papio.
 OC NCBI_TaxID=9557;
 RN [1]
 RE SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Cox L.A., Jett C., Hixson J.E.;
 RI "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
 RI Site Mutation is Associated with Deletion of a Single Exon in a Null
 RI Allele."
 RI Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

```

C D EMBL; AF029692; AAB97887.1; -.
D HSP; P00747; SHPG.
D MEROPS; S01.233; -.
D GO; GO:0005509; F:calcium ion binding; IEA.
D GO; GO:0004263; F:chymotrypsin activity; IEA.
D GO; GO:0008233; F:peptidase activity; IEA.
D GO; GO:0008233; F:peptidase activity; IEA.
D GO; GO:0003809; F:thrombin activity; IEA.
D GO; GO:0004295; F:trypsin activity; IEA.
D GO; GO:0007596; P:blood coagulation; IEA.
D GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
D InterPro; IPR009003; Cys_Ser_trypsin.
D InterPro; IPR001254; Peptidase_S1.
D InterPro; IPR001314; Peptidase_S1A.
D InterPro; IPR003966; Peptidase_S1A_pr.
D Pfam; PF00051; kringle; 1.
D Pfam; PF00089; trypsin; 1.
D PRINTS; PR00722; CHYMOTRYPSIN.
D PRINTS; PR0018; KRINGLE.
D K K Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
F NON TER.
F SEQUENCE 334 AA; 36791 MW; C7DC06E03B965286 CRC64;
E Query Match 21.6%; Score 486.5; DB 6; Length 334;
E Best Local Similarity 36.2%; Pred. No. 1.5e-37;
E Matches 129; Conservative 46; Mismatches 134; Indels 47; Gaps 16;
C 50 CYENGHFYKASTDWMGRPLPNSATVLQTYHAH----RSDALQLGLGKHNYCRNP 105
D 5 CMFNGKRYGKATTVTGPQEWAA-----KEPHSLIFTETTPRAGLEK-NYCRNP 58
C 106 D-NRRRPWCYVQGLKPLVQECMVHDCADGKLKFCQCKTLRPR---FKIIGGFTTIE 161
D 59 DGDVGGPWCYT-TNPRKLYDCVPCASS--SFDGKPKQVCKPCGRVVGCVAAH 115
C 162 QPWFAAYRRHSGSVTVVGGSLISPCWISATHCFIDYPKKEDIVYLGSRNSNTQ 221
D 116 WPMQVSL--RTRFG--MHFCGGTLISPEWVLTAAHCLKSPRPSFYKVLGAHQ----- 165
C 222 GEMKFE--VENLIHKDYSADTLAHHNDIALLKIRSEKGRCAOPSRITQICLPSMNDP 279
D 166 -EYLEPHVQIEIVSKMFSEPA--DIALKLSP-----AITDKVIPACLPSP---P 213
C 280 QF-----GTSCEITGFKENSTDYLPQLKMTVVKLISHRECOOPHYGSEVTTKMLCAA 335
D 214 NYVVADRETCFTGWGTQGT--YGAGLLKEARLPVIEVNCNRYEFLNGRVKASTELCAG 271
C 336 DPQWKTSCQDGGPLVCSLQGRMTLTGIVSGRGCAKDKGQVYTRVSHFLPWI 391
D 272 HLAGTDSQDGGPLVCSLQGRMTLTGIVSGRGCAKDKGQVYTRVSHFLPWI 327
R ILT 28
C 783 PRELIMINARY; PRT; 806 AA.
I O18783
A O18783;
D 01-JAN-1998 (TRENBLrel. 05, Created)
D 01-JAN-1998 (TRENBLrel. 05, last sequence update)
D 01-OCT-2003 (TRENBLrel. 25, last annotation update)
D Plasminogen.
C Macropus eugenii (Tamar wallaby).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Metatheria; Diprotodontia; Macroplidae; Macropus.
OX NCBI_TaxID=9315;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98004511; PubMed=9342350;
RA Lawn R.M., Schwartz K., Patthy L.;
RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog.";
RC Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
D EMBL; AF012297; AAB65760.1; -.
D HSP; P00747; SHPG.
D MEROPS; S01.233; -.
D GO; GO:0005509; F:calcium ion binding; IEA.
D GO; GO:0004263; F:chymotrypsin activity; IEA.
D GO; GO:0008233; F:peptidase activity; IEA.
D GO; GO:0003809; F:thrombin activity; IEA.
D GO; GO:0004295; F:trypsin activity; IEA.
D GO; GO:0007596; P:blood coagulation; IEA.
D GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
D InterPro; IPR009003; Cys_Ser_trypsin.
D InterPro; IPR001254; Peptidase_S1.
D InterPro; IPR001314; Peptidase_S1A.
D InterPro; IPR003966; Peptidase_S1A_pr.
D Pfam; PF00051; kringle; 1.
D Pfam; PF00089; trypsin; 1.
D PRINTS; PR00722; CHYMOTRYPSIN.
D PRINTS; PR0018; KRINGLE.
D K K Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
F NON TER.
F SEQUENCE 806 AA; 90981 MW; 95FAA86DC2064D5 CRC64;
E Query Match 20.8%; Score 469.5; DB 6; Length 806;
E Best Local Similarity 31.9%; Pred. No. 1.8e-35;
E Matches 129; Conservative 43; Mismatches 165; Indels 67; Gaps 15;
QY 13 CLNGGTCVSNKYFSNIHWCNCFKFGQHCIDKSKTCYEGNGHFYRGKASTDTMGRPL 72
DB 448 CSGTGSTVLNAQTTRV-----PSVDTTSHPSD---CMYSGKDYRKGRSTTVTGLCQ 498
QY 73 PWSATVLTQTYHAH---RSDALQLGLGKHNYCRNP-NRRRPWCYVQGLKPLVQECMV 128
DB 499 AWTA-----QBFHRTITFTDTPYPRAGLEENYCRNPDGDPNGWCYT-TNPKLPDYCDI 552
QY 129 HDCADGKLKFCQCKTLRPR---FKIIGGFTTIEQWFAAYRRHSGSVTVVGGSL 185
DB 553 PCQVSPS-SFDGKPRVEPQKPCGRIVGCVCAQPHSEW--QISLRTFGE--HFCGGL 607
QY 186 ISPCWISATHCFIDYPKKEDIVYLGSRNSNTQEMKFEVENLIHKDYSADTLAHH 245
DB 608 IAPQWVLTAAHCLERSQWPGAYKVLG-----LHREVNPEYSQGE 647
QY 246 -----NDIALLKIRSEKGRCAOPSRITQICLPSMNDPQSGTSCETGFKGE 293
DB 648 ICVSLFKGFLAADIALLKL-----NRPAAINDKVIPACLPQDPFWPRTICHVTGWDT 703
QY 294 NSTDYLYPE-OLKMTVVKLISHRECOOPHYGSEVTTKMLCAADPQWKTSCQDGGPL 352
DB 704 QGTS---PRGLLKQASLEFVIDNRVNCNRYEFLNGRVKSTELCAGHLVGRGDSQDGGPL 760
QY 353 VCSLQGRMTLTGIVSGRGCAKDKGQVYTRVSHFLPWI 396

```

761 ICEDDDKVLQGVTSWGLGCRPNKPGYVVRVSRVSWIEDVMK 804

RE Q9 Y91V3 PRELIMINARY; PRT; 868 AA.
AC Q91V3;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Tunicate retinoic acid-inducible modular protease precursor.
TRAMP.

GN Polyandrocarpa misakiensis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Styelidae; Polyandrocarpa.
OX NCBI_TaxID=7723;
RN [1]
RF SEQUENCE FROM N.A.
RC STRAIN=white spot;
RX MEDLINE=99423646; PubMed=10491255;
RT Choshi M., Kawamura K., Fujii N., Yubisui T., Fujiwara S.;
RI "A retinoic acid-inducible modular protease in budding ascidians."
CC Dev. Biol. 214:38-45(1999).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB030007; BAA8252.1; -.
DR HSP; P00763; IDPO.

DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_cysteine.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001134; Peptidase S1.
DR InterPro; IPR001134; Peptidase S1A.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00057; ldl_recept_7; 3.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00530; SRCR; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0261; LDLRECEPTOR.
DR PRINTS; PRO0258; SPRACTRCPTR.
DR SMART; SM00122; LDLr_3.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00202; SR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01209; LDLRA_1; 3.
DR PROSITE; PS00068; LDLRA_2; 3.
DR PROSITE; PS0287; SRCR_2; 2.
DR PROSITE; PS0287; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease; Signal.
FT SIGNAL 1 28 POTENTIAL.
SC SEQUENCE 868 AA; 97660 MW; F71462865F36A6CA CRC64;

ary Match 20.2%; Score 457; DB 5; Length 868;
st Local Similarity 39.2%; Pred. No. 3e-34;
:ches 104; Conservative 43; Mismatches 100; Indels 18; Gaps 9;

QY 139 CCGQKTL----RPRFKIIGBFTTIENQWFAAIVRRHRGSGVTYVCGGSLISPCWISA 194
DE 608 ECGRPVIEAPLPTARIYVGGSGTEPHEWQAGIWL-----PWYVCGGSLIHPCWLTA 662
QY 195 THCFI-DYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKI 253
DE 663 AHCFVREYPIR-DYTIIRIGDHITGVDDTEQLFKIAEIKH-DYNTVT--KENDIALRI 718

QY 254 RSKEGRCAQPSRTIOTICLPSMYNDPQFGTSCITGFGKENSTDY-LYPEOLKMTVVKLI 312
DB 719 ENDARECATIPEVOTVCLPKSSQFDAKTICEVTGKGDATATVAVPVLOAEIPLI 778
QY 313 SHRECOQPHYGYSEVTTXLMCAADPQWKTDCQSGSGGFLVCSLQG--RMILTIGVSWGR 370
DB 779 ANKKCLRDSEY-TQLGPTMFCAGYLTGKGSQSGSGGFLSCRDQSDDRYVVMGIVSMGN 837
QY 371 GCALKDKPGVYTRVSHFLPWIRSH 395
DB 838 GCAKPAFGVYAKVAFIDWIEQMT 862
RESULT 30
Q8N171
ID Q8N171 PRELIMINARY; PRT; 327 AA.
AC Q8N171;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Similar to protease, serine, 8 (Prostasin) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; BC036846; RAH36845.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
SC SEQUENCE 327 AA; 35106 MW; 60458ED817AC1CF3 CRC64;
Query Match 19.7%; Score 445.5; DB 4; Length 327;
Best Local Similarity 38.3%; Pred. No. 1.1e-33;
Matches 105; Conservative 36; Mismatches 100; Indels 33; Gaps 8;
QY 134 GKXKCCGQKTLRPRFKIIGBFTTIENQWFAAIVRRHRGSGVTYVCGGSLISPCWIS 193
DB 69 GRKSAACGQPRMSSR--IVGDRDGRDGEWFWQASL--QHRG---AHVCGGSLIAFWVLT 121
QY 194 ATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKI 253
DB 122 AAHCPRRALPAEYVRVRLGALRLGSTSPRTLSVPVRRVLLPDYSED--GARGDLALLQL 179
QY 254 RSKEGRCAQPSRTIOTICLPSMYNDPQFGTSCITGFGKENSTDYLP-----EQLKM 306
DB 180 R-----RPVLSARVQVCLPVEGARPPPTGTCRVTVGWS-----LRGVPVLPWRPQQ 229
QY 307 TVVKLISRECOQPHYGYSEVTTK-----MLCAADPQWKTDCQSGSGGFLVCSLQGR 359
DB 230 VRVPLDSDTCDGLYHVGADVPQAEIRIVLPGLSLCAGYQGHKDACQSGSGGFLTCLQSGS 289
QY 360 MTLTIGVSWGRCALKDXPGVYTRVSHFLPWIR 393
DB 290 WVLGVVSWGRKCALPFRPGVYTSVATYSPWITQA 323


```

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY190317; AA038062.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001190; Scrv_receptor.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0287; SRCR_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PROTEASE.
DR KX SEQUENCE 558 AA; 60432 MW; 4AC817FCD70D7017 CRC64;
SQ
try Match 19.0%; Score 429; DB 4; Length 558;
at Local Similarity 30.3%; Pred. No. 7.8e-32;
ches 123; Conservative 55; Mismatches 146; Indels 82; Gaps 17;
32 NCP-----KFGGQHCEI-----DKS-KTCYEGNGHYRKGASTDTMGRCLPW 74
198 SCFHAVRCDGVDCVCKLSDELGCVRFDWKSLKLYSGSSHOWLPICSSN-----W 249
75 N---SATVLOQ-----TYAHRSDALQGLGKHNKCRNPNRRRPWCYVQVGLKPL 122
250 NDSYSEKTCQLGPFSAHRTTEVAHRDFANFSILRYN-----ST 289
123 VOECNVH-DCADGK-LKFCQCKTLRPF-KIIGGEFTTIENQWFAAIYRHRGGSVY 179
290 IOESLHRSCHPSQRIYSLQCSHGCLRAWTGRIVGALASDSKWPQVSL---HFG--TTH 344
180 VCGSLISPCWVISATHCFIDYPKK--EDYIVYGRSLNNTQGMKFEVENLILHKDY 237
345 ICGTGLIDAQWVLTAAHCFVTRKVLGKWKVYAGTSLNHLQPEAA---SIABIINSNY 401
238 SADTLAHNDIALLKIRSEGRCAQP---SRTIOTICLPSMYNDPQGTSCETITGK 294
402 TDE--EDDYDIALM-----RLSKPLTSAHIHPACLPMEHQTFSLNETWITGFK 452
295 STDYLYPEQLKMTVVVYKLIHRECOQPHYGVSEVTTKMLCAADPOWKTDCQGD 354
453 ETDDKTSPLFEVQVNLIDFKKNDYLVDSYLTFRMWCAGDLGGRDSCQDGGPLVC 512
355 SLOGRWLTGIVSGRGKALKDKPGVTRVSHFLPWIRSHTKENG 400
513 EQNNRWYLAGVTSWGTGCGQRNKPQVYTKVTEVLPWIYSKMESSAG 558
LT 34
Q9 21
ID Q9BYE1 PRELIMINARY; PRT; 537 AA.
AC Q9BYE1
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mosaic serine protease.
GN MSP8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
EX MEDLINE=21167393; PubMed=11267681;

```

```

RA Kim D.R., Sharmin S., Inoue M., Kido H.;
RT "Cloning and expression of novel mosaic serine proteases with and
RT without a transmembrane domain from human lung.";
RL Biochim. Biophys. Acta 1518:204-209(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB048797; BAB39742.1; -.
DR HSRP; P00763; LDPO.
DR MEROPS; S01.087; -.
DR GO; GO:0016021; C:integral to membrane; NAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR002172; LBL_receptor_A.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001190; Scrv_receptor.
DR Pfam; PF00057; ldl_recept_a; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00192; LDLA; 1.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0287; SRCR_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR HYDROLASE; Protease; Serine protease.
DR KX SEQUENCE 537 AA; 58102 MW; A39FF4E8816DAECF CRC64;
SQ
Query Match 19.0%; Score 428; DB 4; Length 537;
Best Local Similarity 29.7%; Pred. No. 9.2e-32;
Matches 120; Conservative 54; Mismatches 148; Indels 82; Gaps 14;
QY 32 NCP-----KFGGQHCEIDKS-KTCYEGNGHYRKGASTDTMGRCLPW 74
DB 173 SCFHAVRCDGVDCVCKLSDELGCVRFDWKSLKLYSGSSHOWLPICSSN-----W 224
QY 75 NSA-----TVLQQTYYAHRSDALQGLGKHNKCRNPNRRRPWCYVQVGLKPL 122
DB 225 NDSYSEKTCQLGPFSAHRTTEVAHRDFANFSILRYN-----ST 264
QY 123 VOECNVH-DCADGK-LKFCQCKTLRPF-KIIGGEFTTIENQWFAAIYRHRGGSVY 179
DB 265 IOESLHRSCHPSQRIYSLQCSHGCLRAWTGRIVGALASDSKWPQVSL---HFG--TTH 319
QY 180 VCGSLISPCWVISATHCFIDYPKK--EDYIVYGRSLNNTQGMKFEVENLILHKDY 237
DB 320 ICGTGLIDAQWVLTAAHCFVTRKVLGKWKVYAGTSLNHLQPEAA---SIABIINSNY 376
QY 238 SADTLAHNDIALLKIRSEGRCAQP---SRTIOTICLPSMYNDPQGTSCETITGK 294
DB 377 TDE--EDDYDIALM-----RLSKPLTSAHIHPACLPMEHQTFSLNETWITGFK 427
QY 295 STDYLYPEQLKMTVVVYKLIHRECOQPHYGVSEVTTKMLCAADPOWKTDCQGD 354
DB 428 ETDDKTSPLFEVQVNLIDFKKNDYLVDSYLTFRMWCAGDLGGRDSCQDGGPLVC 487
QY 355 SLOGRWLTGIVSGRGKALKDKPGVTRVSHFLPWIRSHTKEE 398
DB 488 EQNNRWYLAGVTSWGTGCGQRNKPQVYTKVTEVLPWIYSKMESE 531
RESULT 35
Q99JUC8
ID Q99JUC8 PRELIMINARY; PRT; 761 AA.
AC Q99JUC8
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neurotrypsin.
GN NT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

```


RA Kim D.R., Sharmin S., Inoue M., Kido H.;
RT "Cloning and expression of novel mosaic serine proteases with and
RF without a transmembrane domain from human lung."
RL Biochim. Biophys. Acta 1518:204-209 (2001).
DE EMBL; AB048796; BAB39741.1; -
DE HSP; P00763; LDPO.
DE MEROPS; S01.087; -
DE GO; GO:0016021; C: integral to membrane; NAS.
DE GO; GO:0006508; P: proteolysis and peptidolysis; NAS.
DE InterPro; IPR009003; Cys_Ser_trypsin.
DE InterPro; IPR002172; LBL_receptor_A.
DE InterPro; IPR001254; Peptidase_S1.
DE InterPro; IPR001314; Peptidase_S1A.
DE InterPro; IPR001190; Sscr_receptor.
DE Pfam; PF00057; ldl_recept_a; 1.
DE Pfam; PF00089; trypsin_1.
DE PRINTS; PR00722; CHYMOTRYPSIN.
DE SMART; SM00192; LDua; 1.
DE SMART; SM00202; SR; 1.
DE SMART; SM00020; Tryp_Spc; 1.
DE PROSITE; PS0287; SRCR_2; 1.
DE PROSITE; PS0240; TRYPSIN_DOM; 1.
DE PROSITE; PS00134; TRYPSIN_HIS; 1.
DE PROSITE; PS00135; TRYPSIN_SER; 1.
DE Hydrolase; Protease, Serine protease.
KW Hydrolyase, Serine protease.
SQ SEQUENCE 581 AA; 62689 MW; 4DABE24D7D5B4A4 CRC64;

3ry Match 18.8%; Score 424; DB 4; Length 581;
st Local Similarity 23.8%; Pred. No. 2.4e-31;
tches 119; Conservative 53; Mismatches 145; Indels 82; Gaps 14;

Qy 32 NCPK-----KFGQHCIEDKS-KTCYEGNGHYRGKASDTMGRCLPW 74
Db 198 SCPHAVRCGVDCVKLSDELGCVRFDWDRKSLIKYSSGSHQWLPICSSN-----W 249
Qy 75 NSA-----TVLQOTYHAHRSDALQLGKIKNYCRNPNRRRPPWCYVQVGLKPL 122
Db 250 NDSYSEKTCRLQGPESAHRTTEVAHRDFANSFSILRYN-----ST 289
Qy 123 VQECM--VHDCADGKLFQCGQKTLRPF-KIIGGEFTTIENQFPAALYRRHGGSVIY 179
Db 290 IQESLHRSHCSQSYIISLQCSHGLRAMTGRIYGGALASDKWFWQVSL---HFG--TTH 344
Qy 180 VCGSLISPCWVISATCFIDYPKK--EDYIVYLGSRSLNSNTQGMKFEVENLILHKDY 237
Db 345 ICGGTLIDAQWVLTAAHCFVTVREKLGKWKYAGTSLNHLQPEAA---SIAEIIINSY 401
Qy 238 SADTLAHHNDIALKIRSGRCAP---SRTIOTICLPSMYNDPQFGTSCBITGFKEN 294
Db 402 TDE--EDDYDIALM-----RLSKPLTSLAHHPACLPMHGQTFSLNETCWTGFKTR 452
Qy 295 STDVLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPOWKTDCQGDGSLVC 354
Db 453 ETDKTSFPLEVQVNLIDFKKNDYLVYDSYLTPTMWCAGDLHGRDSCQGDGSLVC 512
Qy 355 SLOGRMTLTGIVSGRGKALDKPGVYTRVSHFLPWIR 393
Db 513 BQNNRWYLAGVTSWGTGCGQKPKGVYTKVTEVLPWIR 551

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Breast tumor;
RA Strauberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DE EMBL; BC042878; AAH42878.1; -
DE GO; GO:0016020; C: membrane; IEA.
DE GO; GO:0004263; P: chymotrypsin activity; IEA.
DE GO; GO:0008233; P: peptidase activity; IEA.
DE GO; GO:0005044; P: scavenger receptor activity; IEA.
DE GO; GO:0004295; P: trypsin activity; IEA.
DE GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DE InterPro; IPR009003; Cys_Ser_trypsin.
DE InterPro; IPR002172; LBL_receptor_A.
DE InterPro; IPR001254; Peptidase_S1.
DE InterPro; IPR001314; Peptidase_S1A.
DE InterPro; IPR001190; Sscr_receptor.
DE Pfam; PF00057; ldl_recept_a; 1.
DE Pfam; PF00530; SRCR_1; 1.
DE Pfam; PF00089; trypsin_1.
DE PRINTS; PR00722; CHYMOTRYPSIN.
DE SMART; SM00202; SR; 1.
DE SMART; SM00020; Tryp_Spc; 1.
DE PROSITE; PS0287; SRCR_2; 1.
DE PROSITE; PS0240; TRYPSIN_DOM; 1.
DE PROSITE; PS00134; TRYPSIN_HIS; 1.
DE PROSITE; PS00135; TRYPSIN_SER; 1.
KW Protease.
FT NON TER 1
SQ SEQUENCE 471 AA; 52535 MW; ED58CFE6B7C3BCC4 CRC64;

Query Match 18.7%; Score 423; DB 11; Length 471;
Best Local Similarity 32.9%; Pred. No. 2.3e-31;
Matches 114; Conservative 46; Mismatches 133; Indels 54; Gaps 13;

Qy 64 TDTMGRCLPNNSATVLCQTYHAHRSDALQLGKKN-----YCRNPNRRRPPWCYVQ 116
Db 161 TDSK-RTCCQLGFDGSAVYRTEVAHRDITSSFLSEYNTTIOESLYRSCPSRR----YVS 215
Qy 117 VGLKELVQECVHDCADGKLFQCGQKTLRPFKIIGGEFTTIENQFPAALYRRHGG 176
Db 216 L-----QCS-----HGLRAMTGR--IVGALTSESKWQVSL---HFG-- 250
Qy 177 VTYCGSLISPCWVISATCFIDYPKK--EDYIVYLGSRSLNSNTQGMKFEVENILH 234
Db 251 TTHICGGTLIDAQWVLTAAHCFVTVREKLGKWKYAGTSLNHLQPEAA---SISQIIIN 307
Qy 235 KDYSADTLAHHNDIALKIRSGRCAP---SRTIOTICLPSMYNDPQFGTSCBITGFG 291
Db 308 GNYTDE--ODDYDIALI-----RLSKPLTSLAHHPACLPMHGQTFGLNETCWTGFG 358
Qy 292 KENSTDVLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPOWKTDCQGDGSGP 351
Db 359 KTKETDEKTSFPLEVQVNLIDFKKNDYLVYDSYLTPTMWCAGDLHGRDSCQGDGSGP 418
Qy 352 LVCSLQGRMTLTGIVSGRGKALDKPGVYTRVSHFLPWIRSHTKEE 398
Db 419 LVCEQNNRWYLAGVTSWGTGCGQKPKGVYTKVTEVLPWIRKWESE 465

RESULT 39
Q80WN7 PRELIMINARY; PRT; 277 AA.
ID Q80WN7
AC Q80WN7
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Trypsase-6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A.

GN KUK13.
 OE Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22325484; PubMed=12437987;
 RA Olsson A.I., Lundwall A.;
 R1 "Organization and evolution of the glandular kallikrein locus in Mus
 RT musculus.";
 RI Biochem. Biophys. Res. Commun. 299:305-311(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Adams M., Mural R.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY152432; AA078420.1; -.
 DR MGD; MG1:95292; KUK13.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR001254; Peptidase_S1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYSPC; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Hydrolase; Protease; Serine protease.
 KW SEQUENCE 276 AA; 30360 MW; 33E189C67492FDCA CRC64;
 SQ
 Query Match 18.2%; Score 411.5; DB 11; Length 276;
 Best Local Similarity 38.0%; Pred. No. 1.5e-30;
 Matches 98; Conservative 40; Mismatches 81; Indels 39; Gaps 11;
 QY 153 GGEFTTINQWPAIYRRHGGSVTVVCGSLISPCWVISAETHCFIDYPKEDYIVYLG 212
 DE 39 GGYTCLPHSOPWQAALLIRG-----LTCGGVLPKWLTAHC-----RKDGTVHLG 88
 QY 213 RSLRNSNTQGMKEFEVENLIHKDYSADT--LAHNDIALLKIRSKRGCAQPSRTIQT 270
 DE 89 KHALGRVENCEQAMEVVRSPHPEYQVTPTHLNDHDMLELKSP-----VOLSSHVRTL 144
 QY 271 -----CLPSMYNDPQGTSCETGFGKENSVDLYPEQLKMTVKLISHRECCQPHYVG 324
 DE 145 KLSADDCCLPT-----GTCRVSGWGTTTSPQVNYPTKIQOCANIELRDEECRQ--VYP 195
 QY 325 SEVTTKMLCAADPQWKTDSCGGSGPLVCSLQGRMTLTIGIVSWGR--GCALKDKPGVYTR 383
 DE 196 GKITANMLCAGTKGGKDCGGDSGGPLICN--GK--LYGIISWGDFFCQGNRPQVYTR 251
 QY 384 VSHFLPWIRS---HTKEE 398
 DE 252 VSKYLRIWRIIRNTPEQ 269
 RE LT 42
 QY Q8ROP5 PRELIMINARY; PRT; 638 AA.
 AC Q8ROP5;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Kallikrein B. plasma 1.
 GN KUK13.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RT TISSUE=Kidney;

RA Strauberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC026555; AA026555.1; -.
 DR HSPF; P00761; IAN1.
 DR MGD; MG1:102849; KUK13.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0007596; P:blood coagulation; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001177; Apple.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00024; PAN; 4.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00005; APPLEDOMAIN.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00223; APPLE; 4.
 DR SMART; SM00020; TRYSPC; 1.
 DR PROSITE; PS00495; APPLE; 4.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Hydrolase; Protease; Serine protease.
 KW SEQUENCE 638 AA; 71382 MW; CC27C93F4B57C599 CRC64;
 SQ
 Query Match 18.2%; Score 411.5; DB 11; Length 638;
 Best Local Similarity 33.5%; Pred. No. 4.2e-30;
 Matches 94; Conservative 56; Mismatches 106; Indels 25; Gaps 8;
 QY 126 CMVHDCADGKLKFCGGKTLRPFKIIGGFTTINQWPAIYRRHGGSVTVVCGSL 185
 DE 375 CKLVDSPD-----C--TTKINARIVGGTNASLGEPWQVSL--QVKLVSTHLCGSI 423
 QY 186 ISPCWVISAETHCFIDYPKEDYIVYLGSRSLNSNTQGMKEFEVENLIHKDYSADTLAH 245
 DE 424 IGRQWLVTAACHFCGIPYDPVVRVYIGLLSLEITKTPSRIKELIHQEKVS--EGN 481
 QY 246 NDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPQGTSCETGFG--KENSTDIYLYPE 302
 DE 482 YDIALIKLQTP-----LNYTEFQKICLPKSKADTNTVTCWVTGWYTKQGETONI-- 534
 QY 303 QLKMTVKLISHRECCQPHYGVSEVTKMLCAADPQWKTDSCGGSGPLVCSLQGRMTL 362
 DE 535 -LQKATIFLVNPECQK--KYRDYVINKQMICAGYKEGTDACKGDSGGPLVCKHSGRWQL 592
 QY 363 TGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
 DE 593 VGITSMGEGCARKDQPGVYTKVSEYMDWILEKTQSSDVRL 633
 RESULT 43
 Q8CJ16
 ID Q8CJ16 PRELIMINARY; PRT; 371 AA.
 AC Q8CJ16;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Adrenal mitochondrial protease short variant.
 GN AMP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NEDH;
 RA Omer S., Bicknell A.B., Lowry P.J.;
 RT "Identification of a rat adrenal mitochondrial protease.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

```

D EMBL; AF537099; AAN06758.1; ..
D GO; GO:0016020; C:membrane; IEA.
D GO; GO:0004263; F:chymotrypsin activity; IEA.
D GO; GO:0008233; F:peptidase activity; IEA.
D GO; GO:0005044; F:scavenger receptor activity; IEA.
D GO; GO:0004295; F:trypsin activity; IEA.
D GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
D InterPro; IPR009003; Cys_Ser_trypsin.
D InterPro; IPR001254; Peptidase S1.
D InterPro; IPR001314; Peptidase S1A.
D InterPro; IPR001190; Srcr_receptor.
D Pfam; PF00089; trypsin; 1.
D PRINTS; PR00722; CHYMOTRYPSIN.
D SMART; SM00020; Tryp_Spc; 1.
D PROSITE; PS0287; SRCR 2; 1.
D PROSITE; PS0240; TRYPSIN_DOM; 1.
D PROSITE; PS00134; TRYPSIN_HIS; 1.
D PROSITE; PS00135; TRYPSIN_SER; 1.
D K Protease.
D K SEQUENCE 371 AA; 40694 MW; 89A64081D9A1FE26 CRC64;

Query Match 18.2%; Score 411; DB 11; Length 371;
Best Local Similarity 31.7%; Pred. No. 2.4e-30;
Matches 115; Conservative 57; Mismatches 121; Indels 70; Gaps 17;

Q 50 CYEGNGHYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLG---LGKHNKCRPD 106
D :|||:-----WNPALGM-----HICQS-----LGYFRLTQHKAVNLS 82
D 51: CHEG-----

Q 107 ---NRRPWCYVQGLKP---LVQECM---VHDCADGKLK---FQCQKTLRPRFKIIGGE 155
D :|||:-----FAQLSARPSGLVEAWQPSNCPGSRIVSLKCECGARPLASR--LVGQQ 138
D 83 IKLNRSQE---

Q 156 FTTIENQPFPAIYRRHGGSVTVCGGSLISPCWISATHCFIDY--PKKEDIYVYGLR 213
D :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 192
D 139 AVASGRWFOASVMLGSR-----HTCGSVLAPYVWVTAACWYFRLSRLSSWRVHAG- 192

Q 214 SRLNSNT-----QCEMKFEVENILHKDYSADTLAHNDIALIKIRSKGRCQAQPSRTIQ 268
D :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 241
D 193 --LVSHSAVQHOGTM---VEKIIPHLYSAQN--HDYDVALQLRTP-----INFSDTVS 241

Q 269 TICLPSPYNDPQGTSCETIGFKENSTDYLYPEQLKMTVVKLISHRECQQPHYGSEVT 328
D :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 301
D 242 AVCLPAKEHQFPQGSQWVGWGHDPSTHSSDTLQDTWVPLLTDLNCSNCMTSGALT 301

Q 329 TKMLCAADPWKMTDSQCGSGGLVCSLQGRMTLTGIVSWGRCALKDKPQGVYTVVSHFL 388
D :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 361
D 302 HRMLCAGYLDGRADACQSGGGLVCPSGDTHLVGVVSWGRCALPQGVYAKVAEFL 361

Q 389 FWI 391
D :|||:
D 362 DWI 364

```

```

RESULT 45
Q8CDR0 ID Q8CDR0 PRELIMINARY; PRT; 455 AA.
AC Q8CDR0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transmembrane protease.
GN TMPSR85.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;

```


Arakawa T., Hata A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 Hayaishizaki Y.,
 "Functional annotation of a full-length mouse cDNA collection."
 Nature 409:685-690(2001).
 [2]
 SEQUENCE FROM N.A.
 TISSUE=Liver;
 Strausberg R.;
 Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 C -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 D EMBL; AK005346; BAB24114.1; -;
 D ENSP; BC019485; AAI19485.1; -;
 D HSSP; P00750; IRTF.
 D MEROPS; S01.213; -;
 D MGD; MGI:99481; F11.
 D GO; GO:0005576; C:extracellular; IEA.
 D GO; GO:0004263; F:chymotrypsin activity; IEA.
 D GO; GO:0008233; F:peptidase activity; IEA.
 D GO; GO:0004295; F:trypsin activity; IEA.
 D GO; GO:0007596; F:blood coagulation; IEA.
 D GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 D InterPro; IPR001177; Apple.
 D InterPro; IPR009003; Cys_Ser_trypsin.
 D InterPro; IPR003014; PAN_
 D InterPro; IPR001254; Peptidase S1.
 D InterPro; IPR001314; Peptidase_S1A.
 D Pfam; PF00024; PAN; 4.
 D Pfam; PF00089; trypsin; 1.
 D PRINTS; PR00005; APPLIEDOMAIN.
 D PRINTS; PR00722; CHYMOTRYPSIN.
 D SMART; SM00223; APPLE; 4.
 D SMART; SM00020; Tryp_SPC; 1.
 D PROSITE; PS00495; APPLE; 3.
 D PROSITE; PS00240; TRYPSIN_DOM; 1.
 D PROSITE; PS00134; TRYPSIN_HIS; 1.
 D PROSITE; PS00135; TRYPSIN_SER; 1.
 D Hydrolase; Protease; Serine protease.
 K SEQUENCE 624 AA; 69788 MW; 0EEDDEBC56009B97 CRC64;
 S
 Query Match 17.7%; Score 400.5; DB 11; Length 624;
 Best Local Similarity 27.9%; Pred. No. 4.5e-29;
 Matches 126; Conservative 55; Mismatches 144; Indels 127; Gaps 21;
 13 CLNGGTCVSNKYFSNI-----HWC-----NCPKFGGCHCEIDKSK 48
 230 CTHPTCLPTFFPSQAPKESQRLCLKTSGLPSTRITKSHALSGSLQCHRSPV 289
 49 TCVEGNHGYRGKASTDTMG-----RPLPMNSATVLOQTYHAHRSALQ 93
 290 FCHPS---FYN---DTPFLGEELDIVDKQETCQKTCT--NNARQCFPTYF----- 334
 94 LGLGKNYCRPNRRPWCVQVGLKPLVQECNWHDCADG----- 134
 335 -----SHLCLN--ERNRRGRYLKLS-----SNGSPTRILHGRGGTSGVSLRL 375
 135 -KLKFCQCKTLRPRFKIIGEEFTTIENQFWFAAIYRHRGGSVTYCGGSLISPCWVIS 193
 376 CKMDNVCTTK-INPR--VVGGAASVHGEWPWQVTLHI-----SQGLHCGGSLIGNQWILT 427

194 ATHCF--IDYPKKEDYIVYLGSRSLNSNTGEMKFEVENLILHKDYSDTLAHNDIAL 251
 428 AAHCFSGIEPKK--LRYVGIIVNOSINEGTAFRRVQEMIHDQYT--TAESYDIAL 483
 252 KIRSKGRCQAQPSRTIQ-TICLPSMYNDPOFGTSCETGTF-----GKENSTDYLYPEQL 304
 484 KLES-----AMNYTDFQRPICLPSKGRNVAHVTECWVIGWYTGALRGVQST-----L 531
 305 KMTVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDCQGDGSGPLVCSLQGRWTLTG 364
 532 QKAKVPLVSNEEC-QTYRRHKIINMKICAGYEGGKDTCKDGGGGLSKYNGVWHLVG 590
 365 IVSWGRGKALKDKPGVYTRYSHFLPWIRSHTK 396
 591 ITSNGEGCGQKRPDGVYTNVAKYVDMILEKTQ 622
 RESULT 48
 Q72410 PRELIMINARY; PRT; 855 AA.
 ID Q72410;
 AC Q72410;
 DT 01-OCT-2003 (TremBLrel. 25, Created)
 DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
 DE 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE Polysomase 1B protein.
 GN POLYSOMASE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RA MEDLINE=22784742; PubMed=12886014;
 RX Cal S., Lopez-Otin C.;
 RT "An alternative splicing of human polysomase lacking the last serine
 RT protease domain."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:9185-9190(2003).
 DR EMBL; AJ488947; CAD35759.1; -;
 KW Alternative splicing.
 SQ SEQUENCE 855 AA; 91352 MW; 8AF2759D9740CF3F CRC64;
 Query Match 17.7%; Score 400; DB 4; Length 855;
 Best Local Similarity 34.9%; Pred. No. 7.4e-29;
 Matches 98; Conservative 49; Mismatches 110; Indels 24; Gaps 10;
 130 DCADG--KLKFCQCG-QKTLRPRFKIIGEEFTTIENQFWFAAIYRHRGGSVTYCGGSLI 186
 179 DCSGDSFAHCECGLOPAWRMAGRIVGGMEASPGEPFQASL-RENK-----EHFCGAII 233
 187 SPCVVISATHCFIDYPKKEDYIVYLGSRSLNSNTGEMKFEVENLILHKDYSDTLAHN 246
 234 NARWLVSAAHCFNEFPQPTKVAIVGATYLSGSEASTVRAQVQIVKPLYNADTADF-- 291
 247 DIALKIRSKGRCQAQ---SRTIQTICLPSMYNDPOFGTSCETGTCGKENSTDYLYPEQ 303
 252 DVAVLELTS-----PLPFGRIHQVCLPAAATHIFFPSPKCKLISGWYKEDFLVKPEV 344
 304 LKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDCQGDGSGPLVCSLQGRWTLTG 362
 345 LQKATVELLDQALCAS--LYGHSITDRMVCAGYLDGKVDSCQGDGSGPLVCEBFSGRFFL 402
 363 TGIVSWGRGKALKDKPGVYTRYSHFLPMI-RSHTKEENGLA 402
 403 AGIVSWGIGCAARRPGVYARVTRLDWILEATTKASMPLEA 443
 RESULT 49
 ID Q72411 PRELIMINARY; PRT; 1059 AA.
 AC Q72411;
 DT 01-OCT-2003 (TremBLrel. 25, Created)
 DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Polysaccharide 1A protein.
 GN POLYSACCHARIDE
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA MEDLINE=22784742; PubMed=12886014;
 RX Cal S., Lopez-Otin C.;
 RI "An alternative splicing of human polysaccharide lacking the last serine
 RT protease domain."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:9185-9190(2003).
 DR EMBL; AJ488946; CAB35758.1; -
 KA Alternative splicing.
 SQ SEQUENCE 1059 AA; 114020 MW; 17D27A2D99F2A264 CRC64;
 ary Match 17.7%; Score 400; DB 4; Length 1059;
 at Local Similarity 34.9%; Pred. No. 9.6e-29;
 tches 98; Conservative 49; Mismatches 110; Indels 24; Gaps 10;
 Qy 130 CDADG--KLKPCQG-QXTLPRFKLIGBETTTIENQWPAIYRHRGGSTVYVCGSLI 186
 Db 179 DCSGDSAEHCECEGLOPAWMAGRIVGMEASPEFPWQASL-RENK----EHFCGAII 233
 Qy 187 SPQVIGATHCFIDYPKKEDYIVYGRSLNSNTQGMKPEVENLILHKDYSADTLAHN 246
 De 234 NARLVSAHCFNEFQPTKVAIVGATYLSGEASTVRAQVIVGHPLYNADTF-- 291
 Qy 247 DIALLKTRSEKGRCAQP----SRTQITCLPSMNDPOFGTSCITGPKENSTDIYLPQ 303
 De 292 DVAVLELTS-----PLPGRHIQVCLPAATHIFPPSKKCIISGMYLKEDFLVKPEV 344
 Qy 304 LKMTWVWLKSHRECCQPHYGVSEVTTKMLCAADPQWKTDCQDGGSGPLVC-SLQGRMTL 362
 Db 345 LQATVELLDALCAS--LYGHSITDRMVCAGYLDGVDSQDGGSGPLVCEPSPGRFL 402
 Qy 363 TGIVSWRGCGKALKDKPGVYTVSHFLPMI-RSHYKEENGIA 402
 Db 403 AGIVSWGIGCAEARPPGVYARVTRLDWILEATTAKSMPLA 443
 RE 50
 IT 17
 AC Q9JUI7 PRELIMINARY; PRT; 855 AA.
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Membrane bound serine protease (Membrane bound arginine specific
 DE serine protease).
 GN MBSP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Jejunum;
 RA Tsuzuki S.;
 RI "A membrane bound serine protease expressed in rat small intestine."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Duoenum;
 RA Inoue H., Takahashi K., Kishi K.;
 RI "membrane-bound arginine specific serine protease."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 DR EMBL; AB037898; BAB03502.1; -

DR EMBL; AB049189; BAB13765.1; -
 DR PIR; JC7731; JC7731.
 DR HSRP; P00763; LDPO.
 DR MEROPS; S01.302; -
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0002233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000859; CUB; 2.
 DR InterPro; IPR003003; Cys_Ser_trypsin.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00431; CUB; 2.
 DR Pfam; PF00057; ldl_recept_a; 4.
 DR Pfam; PF00089; trypsin_1_1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR SMART; SM00042; CUB; 2.
 DR SMART; SM00192; LDLA; 3.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS01180; CUB; 2.
 DR PROSITE; PS01209; LDLRA_1; 2.
 DR PROSITE; PS00068; LDLRA_2; 4.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 FT VARIANT 665 665 K -> N.
 SQ SEQUENCE 855 AA; 94955 MW; 35806B7ECF6CF03D CRC64;
 Query Match 17.6%; Score 397; DB 11; Length 855;
 Best Local Similarity 28.2%; Pred. No. 1.4e-28;
 Matches 120; Conservative 69; Mismatches 153; Indels 84; Gaps 22;
 QY 1 SNELHQVPSNCDCLNGGTCVSNKYPSNIHW-----CNCPRKKFGQHCE- 43
 Db 481 SPDRH--CRCNATHQFMC-KNQFCKPLFWVCDSDVNDGCGSDEEGSCPA--GSFKCSN 534
 QY 44 ----IDSKTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYAHRS DALQLGKHN 100
 Db 535 GKCLPQSOQC---NGKDCGDSDEA---SCDNVNAVSTKYTYRC-----QNGL----- 578
 QY 101 YCRNPDRRRPWCYVQVGLKPLVQECMVHDCADG--KLKFCQCGQKTLPRFKIIGGEFTT 158
 Db 579 -CLNKGK---PEC---DGKK-----DCSDGSDKNCDCGLRSFTKQARVVGTTNAD 622
 QY 159 IENQWPAIYRHRGGSTVYVCGSLISPCWVISATHCFID-----YPKEDYIVYLG- 212
 Db 623 EGEPWQVSLHALGQG---HLCGASLISPDWLVSAHCFQDETIFKYSDDHMTMTAFGL 678
 QY 213 RSRLNSNTQGMKPEVENLILHKDYSADTLAHNDIALKIRSKEGRCQAQPSRTIQTICL 272
 Db 679 LQSKRSASGVQEHKXLIITHPSENFDFDY--DIALLELE---KPAEYSTVVRPCL 732
 QY 273 PSMYNDPQGTSCETGFG--KENSTDYVYBPQKQVTVKLVSHRECOOPHYGSEVTK 330
 Db 733 PNTHTVFPAGKAIWVTGHTKEGCTGALI---LQKEIRVINQTTCE--LLPQOIIPR 787
 QY 331 MLCADDPQWKTDCQDGGPL-VCSLOGRMTLTGIVSWRGCGKALKDKPGVYTVSHFLP 389
 Db 788 MMCVGLSGVDSQDGGSGPLSSVEKGRIFQAGVSWRGCGCAQRNPGVYTRIPEDVR 847
 QY 390 WIRSHT 395
 Db 848 WIKQET 853

Search completed: May 25, 2004, 14:57:17
 Job time : 69.7546 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: May 25, 2004, 14:43:00 ; Search time 98.5425 Seconds
(without alignments)

1155,508 Million cell updates/sec

le: US-09-880-503-6

fect score: 2257
uence: 1: SNELHQPNSCDLNGTCV.....VSHFLPWKSHYKENGIAL 403

ring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

rchd: 1586107 segs, 282547505 residues

al number of hits satisfying chosen parameters: 1586107

imum DB seq length: 0

imum DB seq length: 200000000

t-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

abase : A Geneseq_29Jan04.*

1: geneseqp1980s.*

2: geneseqp1990s.*

3: geneseqp2000s.*

4: geneseqp2001s.*

5: geneseqp2002s.*

6: geneseqp2003as.*

7: geneseqp2003bs.*

8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Full No.	Score	Query Match	Length	DB ID	Description
1	2257	100.0	403	5	Aae16547 Human uro
2	2243	99.4	411	1	AAP50871 Sequence
3	2243	99.4	411	2	AAR06244 Urokinase
4	2243	99.4	411	2	AAR62991 Pro-uroki
5	2243	99.4	411	3	AAY92836 Urokinase
6	2243	99.4	411	5	Aae16544 Human uro
7	2243	99.4	431	1	AAP50114 Sequence
8	2243	99.4	431	1	AAP60783 Human uro
9	2243	99.4	431	1	AAP70258 Sequence
10	2243	99.4	431	1	AAP80430 Deduced A
11	2243	99.4	431	1	AAP81204 Pro-uroki
12	2243	99.4	431	1	AAP92119 Natural h
13	2243	99.4	431	2	AAR07112 Human pro
14	2243	99.4	431	2	AAR04253 Human pro
15	2243	99.4	431	2	AAR63141 Full leng
16	2243	99.4	431	2	AAR47903 Pro-uroki
17	2243	99.4	431	3	AAY50869 Human uro
18	2243	99.4	431	3	AAY99591 Human pla
19	2243	99.4	431	4	AAB84605 Amino aci
20	2243	99.4	431	5	AAE17128 Human uPA
21	2243	99.4	431	5	AAY92228 Human pla
22	2243	99.4	431	5	AAG79460 sc-uPA. 1
23	2243	99.4	431	6	ABR55855 Human uro
24	2243	99.4	431	6	ABU56547 Lung canc
25	2243	99.4	431	6	ABU56708 Lung canc

ALIGNMENTS

RESULT 1

AAE16547

ID AAE16547 standard; protein; 403 AA.

XX

AC AAE16547;

XX

DT 09-APR-2002 (first entry)

XX

DE Human urokinase-type plasminogen activator scuPA delta136-143 mutant.

XX

KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
clotting disorder; uterine contraction disorder; respiratory disease;
male impotence; adult respiratory distress syndrome; scuPA delta136-143;
single chain urokinase; mutant; mutein.

XX

26	2243	99.4	431	6	ABU11076
27	2243	99.4	431	6	ABR92137 Human cer
28	2243	99.4	434	2	AAR20537 Amidated
29	2243	99.4	434	2	AAR20538 Amidated
30	2243	99.4	436	2	AAR20536 Amidated
31	2241	99.3	431	7	ADe25745 Human pro
32	2240	99.2	430	2	AAW24578 Inhibitor
33	2240	99.2	431	1	AAp91886 Sequence
34	2240	99.2	431	1	AAp94764 Non-glyco
35	2240	99.2	431	6	AAE37128 Human uro
36	2240	99.2	431	7	ADD46429 Human pro
37	2239	99.2	411	2	AAr34584 Mutant hu
38	2239	99.2	431	5	AAU99230 Human pla
39	2238	99.2	411	1	AAp96146 Sequence
40	2238	99.2	411	1	AAp99226 Pro-uroki
41	2238	99.2	411	4	AAAB74797 Prourokin
42	2238	99.2	431	5	AAU99229 Human pla
43	2237	99.1	411	2	AAr62999 Pro-uroki
44	2237	99.1	411	2	AAr62992 Pro-uroki
45	2237	99.1	411	2	AAr62998 Pro-uroki
46	2237	99.1	411	2	AAr62993 Pro-uroki
47	2235	99.0	411	2	AAr63000 Pro-uroki
48	2235	99.0	411	2	AAy39343 Human pro
49	2235	99.0	411	2	AAy42284 Human pro
50	2235	99.0	411	4	AAAB20489 Human pro
51	2235	99.0	412	2	AAr10334 Recombina
52	2235	99.0	431	5	AAU99236 Human pla
53	2234	99.0	411	2	AAr07904 Human pro
54	2234	99.0	411	2	AAr07902 Human pro
55	2234	99.0	411	2	AAr07903 Human pro
56	2234	99.0	411	2	AAr40225 PUK. 2/19
57	2234	99.0	431	1	AAp30041 Sequence
58	2234	99.0	431	5	AAU99231 Human pla
59	2233	98.9	411	2	AAr62997 Pro-uroki
60	2233	98.9	411	7	ADe85377 Human sin
61	2233	98.9	431	5	AAU99232 Human pla
62	2232	98.9	431	1	AAp70250 Sequence
63	2231	98.8	411	2	AAW13634 Human rat
64	2231	98.8	411	2	AAr47958 PUK S26T.
65	2231	98.8	411	2	AAr62995 Pro-uroki
66	2231	98.8	411	2	AAr62996 Pro-uroki
67	2231	98.8	431	5	AAU99238 Human pla
68	2230	98.8	411	1	AAp91684 Sequence
69	2230	98.8	411	2	AAr10056 Pro-uroki
70	2230	98.8	411	5	AAE13269 Human pro
71	2230	98.8	431	5	AAU99237 Human pla
72	2229	98.8	411	2	AAr10057 Pro-uroki
73	2229	98.8	411	2	AAr62994 Pro-uroki
74	2229	98.8	411	2	AAr63003 Pro-uroki
75	2229	98.8	411	2	AAr63001 Pro-uroki

Oe	Homo sapiens.
Oe	Synthetic.
Xx	WO200197752-A2.
Xf	
Xy	27-DEC-2001.
Xz	
X1	13-JUN-2001; 2001WO-US018976.
X2	
X3	20-JUN-2000; 2000US-0212874P.
X4	(TYPE-) UNIV PENNSYLVANIA.
X5	Cines DB, Higazi AA;
X6	WPI; 2002-122240/16.
X7	N-PSDB; RAD27080.
X8	
X9	Composition for modulating muscle cell and tissue contractility for treating atherosclerosis, asthma, hypertension, glaucoma, impotence, comprising domains from urokinase-type plasminogen activator.
Xa	
Xb	Claim 22; Fig 1F; 117pp; English.
Xc	
Xd	The invention relates to a composition comprising one or more domains of urokinase-type plasminogen activator (uPA). The composition is used to modulate the contractility and angiogenic activity of a mammalian muscle, endothelial cell or tissue. The composition is used for treating stroke, hypotension, hypertension, atherosclerosis, heart attack, microvascular occlusions, thrombotic microangiopathies, surgically induced thrombotic disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma, diabetic retinopathy, wound healing, clotting disorder, uterine contraction disorder, male impotence, respiratory disease or condition such as asthma, adult respiratory distress syndrome, primary pulmonary hypertension, microvascular thrombotic occlusion, and a disorder associated with chronic intrapulmonary fibrin formation. The present sequence is human urokinase-type plasminogen activator (uPA) single chain urokinase (scuPA) deletion mutant designated as scuPA deltail36-143 Sequence 403 AA;
Xe	
Xf	ary Match 100.0%; Score 2257; DB 5; Length 403;
Xg	st Local Similarity 100.0%; Pred. No. 7.8e-176; Indels 0; Gaps 0;
Xh	ches 403; Conservative 0; Mismatches 0;
Xi	
Xj	1 SNEIHOVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGRGHFYRG 60
Xk	1 SNEIHOVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGRGHFYRG 60
Xl	
Xm	61 KASITDTMGRCPLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPCYYOVGLK 120
Xn	
Xo	61 KASITDTMGRCPLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPCYYOVGLK 120
Xp	
Xq	121 PLVQCWMVHDCADGKLKFCQGOKTLRPRFKIIGGEFTTIENOPFAAIYRRHRGGSVTV 180
Xr	121 PLVQCWMVHDCADGKLKFCQGOKTLRPRFKIIGGEFTTIENOPFAAIYRRHRGGSVTV 180
Xs	
Xt	181 CGGSILISPCWLVSATHCFIDYPKEDYIVLGRSLNSNTQGMKFVENLLHKDYSAD 240
Xu	181 CGGSILISPCWLVSATHCFIDYPKEDYIVLGRSLNSNTQGMKFVENLLHKDYSAD 240
Xv	
Xw	241 TLIAHNDIALLKIRKEGCAQPSRTIQICLPMSMYNDPQGTSCBITGFGENSTDIYL 300
Xx	241 TLIAHNDIALLKIRKEGCAQPSRTIQICLPMSMYNDPQGTSCBITGFGENSTDIYL 300
Xy	
Xz	301 PSOLKMTVKLLSHRECCQPHYVGSEVTVMKLCADDPQWKTSQQGDSSGPLVCSLGSRM 360
Ya	301 PSOLKMTVKLLSHRECCQPHYVGSEVTVMKLCADDPQWKTSQQGDSSGPLVCSLGSRM 360
Yb	
Yc	361 TLTGVSWGRGCALDKDPGYTRVSHFLPWIRSHTKTEENGLAL 403
Yd	361 TLTGVSWGRGCALDKDPGYTRVSHFLPWIRSHTKTEENGLAL 403

RESULT 2	
AAP50871	AAP50871 standard; protein; 411 AA.
XX ID	XX AAP50871 standard; protein; 411 AA.
XX AC	XX AAP50871;
XX	
XX	30-NOV-1991 (first entry)
XX	
XX	Sequence encoded by cDNA sequence for human urokinase zymogen (Japanese Patent Application No.37119/84).
DE DE	
XX	Thrombolytic agent; plasminogen activator activity; fibrin affinity;
XX	enzyme.
XX	
XX	Homo sapiens.
OS	
XX	
XX	Key
PH	Location/Qualifiers
FT	Disulfide-bond 50..131
FT	Disulfide-bond 71..113
FT	Disulfide-bond 102..126
FT	Disulfide-bond 148..279
FT	Cleavage-site 158..159
FT	/note="potential cleavage site which generates the two-chain form from the zymogen"
FT	189..205
FT	Disulfide-bond 197..268
FT	Disulfide-bond 293..362
FT	Disulfide-bond 325..341
FT	Disulfide-bond 352..380
XX	
XX	EP139447-A.
PN	
XX	
XX	02-MAY-1985.
PD	
XX	
XX	07-SEP-1984; 84EP-00306117.
PF	
XX	
PR	13-SEP-1983; 83JP-00170354.
PR	17-OCT-1983; 83JP-00195051.
XX	
XX	(GREC) GREEN CROSS CORP.
PA	
XX	
XX	Kasai S, Arimura H, Mori K, Suyama T;
PI	
XX	WPI; 1985-106530/18.
DR	
XX	
XX	New urokinase zymogen - useful as thrombolytic agent.
PT	
XX	
PS	Disclosure; Page 12; 30pp; English.
XX	
XX	Zymogen AAP50871 is the inactive precursor form of human urokinase.
CC	
CC	Urokinase zymogen is cleaved into the two-chain form composed of
CC	characteristic urokinase H (molecular wt. of 30,000) and L (molecular
CC	wt. of 20,000) chains when treated with catalytic amounts of plasmin. The
CC	patentors claim a new urokinase zymogen which has mol. wt. ca. 50,000, a
CC	single chain molecular structure, and selective affinity for fibrin. It
CC	is a thrombolytic agent which manifests its plasminogen activator
CC	activity on cleavage by proteolytic enzymes (e.g. plasmin) and has higher
CC	affinity for fibrin than known forms of urokinase
XX	
XX	Sequence 411 AA;
SQ	
	Query Match 99.4%; Score 2243; DB 1; Length 411;
	Best Local Similarity 98.1%; Pred. No. 1.1e-174;
	Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1
QY	1 SNEHLOVPSNDCNLNGGTCVSNKYPSNTHWCNCPKKGQHQCEIDKSKTCYEGNGHFYRG 60
Db	1 SNEHLOVPSNDCNLNGGTCVSNKYPSNTHWCNCPKKGQHQCEIDKSKTCYEGNGHFYRG 60
QY	61 KASTDTMGRCPLPWNSATVLQCTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYQVQGLK 120

1 61 KASDTMGRPCLPNSATVLAQTYHAHRS DALQLGLGKHN YCRNPDNRRPWCYVQVGLK 120
121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
121 PLVQECMVHDCADGKSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRH 180
173 RGGSVTVVCGSLSPCWVISAHCFTIDYPKKEDYIVLGRSLNNTQGMKFEVENLI 232
181 RGGSVTVVCGSLSPCWVISAHCFTIDYPKKEDYIVLGRSLNNTQGMKFEVENLI 240
233 LHKDYSADTLAHHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 292
241 LHKDYSADTLAHHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 300
293 ENSDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPWKTDSCQDSDGGPL 352
301 ENSDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPWKTDSCQDSDGGPL 360
353 VCSLQGRMTLTGIVSWGRGCALCKDQPGVYTRVSHFLPWIRSHTKENGLAL 403
361 VCSLQGRMTLTGIVSWGRGCALCKDQPGVYTRVSHFLPWIRSHTKENGLAL 411

ULT 3
06244
AAR06244 standard; protein; 411 AA.

AAR06244;
07-DEC-1990 (first entry)
Urokinase precursor protein.
Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis;
myocardial infarction.
Homo sapiens.
EP380334-A.
01-AUG-1990.
25-JAN-1990; 90EP-00300772.
27-JAN-1989; 89JP-00016406.
17-MAY-1989; 89JP-00121405.
(GREG) GREEN CROSS CORP.
Matsuda H, Ueda Y, Tamanouchi K;
WPI; 1990-233117/31.
Urokinase precursor-lipid composite - used as thrombolytic agent, having
prolonged half-life in the blood, enhanced bioavailability and improved
activity.
Claim 3; Fig 1; 11pp; English.
By forming a precursor-lipid composite, the half-life of this
thrombolytic agent in the blood may be increased, exhibiting improved
activity without abnormal acceleration of fibrinolytic activity. Compound
is useful as a thrombolytic agent in treatment of cerebral thrombosis,
myocardial infarction etc
Sequence 411 AA;
Query Match 99.4%; Score 2243; DB 2; Length 411;
Best Local Similarity 99.1%; Pred. No. 1.1e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
1 SNEHQVPSNCDLNGTGVSNKYFNSIHMNCNPKFGGQHCEIDKSKTCYEGNGHFYRG 60

Db 1 SNEHQVPSNCDLNGTGVSNKYFNSIHMNCNPKFGGQHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASDTMGRPCLPNSATVLAQTYHAHRS DALQLGLGKHN YCRNPDNRRPWCYVQVGLK 120
Db 61 KASDTMGRPCLPNSATVLAQTYHAHRS DALQLGLGKHN YCRNPDNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
Db 121 PLVQECMVHDCADGKSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRH 180
QY 173 RGGSVTVVCGSLSPCWVISAHCFTIDYPKKEDYIVLGRSLNNTQGMKFEVENLI 232
Db 181 RGGSVTVVCGSLSPCWVISAHCFTIDYPKKEDYIVLGRSLNNTQGMKFEVENLI 240
QY 233 LHKDYSADTLAHHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 292
Db 241 LHKDYSADTLAHHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 300
QY 293 ENSDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPWKTDSCQDSDGGPL 352
Db 301 ENSDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPWKTDSCQDSDGGPL 360
QY 353 VCSLQGRMTLTGIVSWGRGCALCKDQPGVYTRVSHFLPWIRSHTKENGLAL 403
Db 361 VCSLQGRMTLTGIVSWGRGCALCKDQPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 4
AAR62991
ID AAR62991 standard; protein; 411 AA.
XX AC AAR62991;
XX DT 25-MAR-2003 (revised)
DT 21-SEP-1995 (first entry)
XX DE Pro-urokinase.
XX KW Pro-urokinase; thrombolysis; fibrin clot lysis.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Domain 297..313 /note= "flexible loop"
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX 12-JAN-1995.
XX PD 28-JUN-1994; 94WO-US007278.
XX PF 02-JUL-1993; 93US-00087163.
XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX PI Liu J, Gurewich V;
XX WPI; 1995-060991/08.
XX Pro-urokinase mutants - have thrombolytic activity but reduced

fibrinogenolysis activity and non-specific plasminogen activation.
Disclosure; Fig 1; 46pp; English.
AAR62991 is the wild type pro-urokinase, from which the new mutants described in AAR62992-R63008 were derived. These mutants retain the thrombolytic activity of the wild type protein, useful for the treatment of thromboembolism, but have a reduced fibrinogenolysis activity and non-specific plasminogen activation. The mutants can therefore be used for the lysis of fibrin clots without inducing systemic bleeding, as can be the case with the wild type protein. (Updated on 25-MAR-2003 to correct PN field.)
Sequence 411 AA;
ery Match 99.4%; Score 2243; DB 2; Length 411;
st Local Similarity 98.1%; Pred. No. 1.1e-174;
tches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHYRG 60
DI 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHYRG 60
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRPPWCYVQVGLK 120
DI 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK-----LKFCQCKTLRPRFKIIGGEFTTIENQFPFAAIYRRH 172
DI 121 PLVQECMVHDCADGKPPPEELKFCQCKTLRPRFKIIGGEFTTIENQFPFAAIYRRH 180
QY 173 RGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVYVLSRLNSNTQGEKFEVENLI 232
DI 181 RGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVYVLSRLNSNTQGEKFEVENLI 240
QY 233 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCBITGFK 292
DI 241 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCBITGFK 300
QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYTGSEVTTKMLCAADPOWKTDSQGDGSGPL 352
DI 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYTGSEVTTKMLCAADPOWKTDSQGDGSGPL 360
QY 353 VCSLQGRMTLTIGVSGRGCAKDKPGYTVTRVSHFLPWIRSHTEENGLAL 403
DI 361 VCSLQGRMTLTIGVSGRGCAKDKPGYTVTRVSHFLPWIRSHTEENGLAL 411
RI LT 5
AJ 2836
II AAY92836 standard; protein; 411 AA.
AC AAY92836;
XZ 29-AUG-2000 (first entry)
XZ Urokinase plasminogen activator (uPA).
XZ N-terminal; pro-uPA; urokinase plasminogen activator; receptor; uPAR;
XZ anti-cancer; anti-metastatic; anti-proliferative; anti-atherosclerotic;
XZ anti-thrombotic; anti-angiogenic; anti-inflammatory; anti-arthritis;
XZ anti-fibrotic; apoptotic; vasotropic; anti-diabetic; ophthalmological;
XZ thrombolytic.
XZ Homo sapiens.
XZ Key Location/Qualifiers
DI Disulfide-bond 11..19
DI Disulfide-bond 13..31
DI Disulfide-bond 33..42
DI Disulfide-bond 50..131
DI Disulfide-bond 71..113
DI Disulfide-bond 102..126

FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO200026353-A1.
XX 11-MAY-2000.
XX 28-OCT-1999; 99WO-US025210.
XX 29-OCT-1999; 98US-00181816.
XX (ANGS-) ANGSTROM PHARM INC.
XX Mazar AP, Jones TR;
XX PI
XX PS
XX DR
XX WIPI; 2000-365605/31.
XX New cyclic peptide, useful for treatment or diagnosis of e.g. tumors and other diseases involving cell proliferation or migration, targets the urokinase plasminogen activator receptor.
XX Disclosure; Fig 1; 93pp; English.
CC The present sequence shows the wild-type urokinase plasminogen activator (uPA). Cyclic peptides based on the amino acids residues 20-30 (the receptor-binding region) of uPA are claimed. These cyclic peptides target the uPA receptor (uPAR), allowing therapeutic or diagnostic agents to be delivered to uPAR-expressing cells. The cyclic peptides are used, optionally when linked to a therapeutic agent, to inhibit migration, invasion and proliferation of cells, or angiogenesis, or to induce apoptosis. Particularly they are used, in human or veterinary medicine, to treat diseases characterized by these processes, e.g. solid tumors, leukaemia or lymphoma (or their metastases); benign hyperplasia; atherosclerosis; restenosis; ischaemia; deep vein thrombosis; neovascular glaucoma; diabetic retinopathy; arthritis; fibrosis; bone fracture etc., most particularly growth, invasion and metastasis of tumors. When labeled, the cyclic peptides can be used for diagnostic detection of uPAR (a marker of metastasis) on cells, tissues etc., in vivo or in vitro, and when immobilized they are used to isolate uPAR or cells that express them. The cyclic peptides are stable, soluble in water, bind strongly to uPAR. The cyclic peptides are inexpensive to produce and may be derivatized to attachment of therapeutic or diagnostic agents without significantly affecting their binding. Since they target uPAR, they should have relatively low systemic toxicity and only low doses are required
XX Sequence 411 AA;
Query Match 99.4%; Score 2243; DB 3; Length 411;
Best Local Similarity 98.1%; Pred. No. 1.1e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHYRG 60
DI 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHYRG 60
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRPPWCYVQVGLK 120
DI 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK-----LKFCQCKTLRPRFKIIGGEFTTIENQFPFAAIYRRH 172
DI 121 PLVQECMVHDCADGKPPPEELKFCQCKTLRPRFKIIGGEFTTIENQFPFAAIYRRH 180
QY 173 RGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVYVLSRLNSNTQGEKFEVENLI 232
DI 181 RGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVYVLSRLNSNTQGEKFEVENLI 240
QY 233 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCBITGFK 292
DI 241 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCBITGFK 292

D 241 LHKDYSADTLAHNDIALALKIRSKGRCAQPSRTIQTICLPSPMYNDPQFTSCIEITGFGK 300
C 293 ENSTDYLYPEOLKMTVVKLIHRECOQPHYVGVSEVTTKMLCAADPOWKTDSCQDGGGGL 352
F 301 ENSTDYLYPEOLKMTVVKLIHRECOQPHYVGVSEVTTKMLCAADPOWKTDSCQDGGGGL 360
C 353 VCSLQGRMTLTGIVSWRGCGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
D 361 VCSLQGRMTLTGIVSWRGCGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
F ULT 6
F 16544
F AAEL16544 standard; protein; 411 AA.
F AAEL16544;
F 09-APR-2002 (first entry)
F Human urokinase-type plasminogen activator tcuPA and scuPA protein.
F Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
F stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
F microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
F tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
F clotting disorder; uterine contraction disorder; respiratory disease;
F male impotence; adult respiratory distress syndrome; tcuPA; scuPA;
F two chain urokinase; single chain urokinase.
C Homo sapiens;
C WO2001:197752-A2.
F 27-DEC-2001.
F 13-JUN-2001; 2001WO-US018976.
F 20-JUN-2000; 2000US-0212874P.
F (UYPE-) UNIV PENNSYLVANIA.
F Cines DB, Higazi AA;
F WPI; 2002-122240/16.
F N-PSDB; AAD27077.
F Composition for modulating muscle cell and tissue contractility for
F treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
F comprising domains from urokinase-type plasminogen activator.
F Claim 9; Fig 1C; 117pp; English.
F The invention relates to a composition comprising one or more domains of
F urokinase-type plasminogen activator (uPA). The composition is used to
F modulate the contractility and angiogenic activity of a mammalian muscle,
F endothelial cell or tissue. The composition is used for treating stroke,
F hypotension, hypertension, atherosclerosis, heart attack, microvascular
F occlusions, thrombotic microangiopathies, surgically induced thrombotic
F disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
F invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
F diabetic retinopathy, wound healing, clotting disorder, uterine
F contraction disorder, male impotence, respiratory disease or condition
F such as asthma, adult respiratory distress syndrome, primary pulmonary
F hypertension, microvascular thrombotic occlusion, and a disorder
F associated with chronic intrapulmonary fibrin formation. The present
F sequence is human urokinase-type plasminogen activator (uPA) two chain
F urokinase (tcuPA) and single chain urokinase (scuPA) protein
F Sequence 411 AA;
F Query Match 99.4%; Score 2243; DB 5; Length 411;
F Local Similarity 98.1%; Pred. No. 1.1e-174;
F atches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKFGQHQCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKFGQHQCEIDKSKTCYEGNGHFYRG 60
QY 61 KASDTMGRPCLPWNSATVLOOTYHAHRSDALQGLGKHNYCRNPDNERRRWCYVQVGLK 120
DB 61 KASDTMGRPCLPWNSATVLOOTYHAHRSDALQGLGKHNYCRNPDNERRRWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK-----LKFCQGGKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
DB 121 PLVQECMVHDCADGKPPPEELKFCQGGKTLRPRFKIIGGEFTTIENQWFAAIYRRH 180
QY 173 RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGEKKEVENLI 232
DB 181 RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGEKKEVENLI 240
QY 233 LHKDYSADTLAHNDIALALKIRSKGRCAQPSRTIQTICLPSPMYNDPQFTSCIEITGFGK 292
DB 241 LHKDYSADTLAHNDIALALKIRSKGRCAQPSRTIQTICLPSPMYNDPQFTSCIEITGFGK 300
QY 293 ENSTDYLYPEOLKMTVVKLIHRECOQPHYVGVSEVTTKMLCAADPOWKTDSCQDGGGGL 352
DB 301 ENSTDYLYPEOLKMTVVKLIHRECOQPHYVGVSEVTTKMLCAADPOWKTDSCQDGGGGL 360
QY 353 VCSLQGRMTLTGIVSWRGCGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
DB 361 VCSLQGRMTLTGIVSWRGCGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 7

AAP50114

ID AAP50114 standard; protein; 431 AA.

AC AAP50114;

XX 27-SEP-1991 (first entry)

XX Sequence encoded by the signal sequence and noncoding region of the pro-

DE UK structural gene (Sequence II).

XX Enzyme; thrombosis therapy; embolic disease; single-chain pro-urokinase.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..20

FT Domain 21..177

FT Domain 179..431

FT Domain /label= A chain

FT Domain /label= B chain

XX EP154272-A.

XX 11-SEP-1985.

XX 23-FEB-1985; 85EP-00102031.

XX 27-FEB-1984; 84JP-00037119.

XX 31-JAN-1985; 85JP-00017969.

XX (GREC) GREEN CROSS CORP.

XX Hiramatsu R, Kaneda T, Nagai M, Rimura H, Nishida M, Suyama T;

XX WPI; 1985-224693/37.

XX N-PSDB; AAN50138.

XX Glycosylated single-chain pro-urokinase - prepd. by cultivating animal

XX cells transformed by DNA prepd. from m RNA.

XX Disclosure; Page 8-10; 64pp; English.

F	EP231883-A.		XX	AAP80430;	
X	12-AUG-1987.		AC		
X	29-JAN-1987;	87EP-00101209.	DT	25-MAR-2003 (revised)	
X	31-JAN-1986;	86JP-00017734.	DT	14-SEP-1990 (first entry)	
P	30-JAN-1987;	87JP-00018626.	XX	Deduced AA sequence of the single chain urokinase plasminogen activator (SCU-PA) cDNA insert prepared from human Hep3 cells.	
X	(SAGA) SAGAMI CHEM RES CENTRE.		DE	Single chain urokinase plasminogen activator (SCU-PA); human Hep3 cells;	
P	(NIPS) NIPPON SODA CO.		XX	glyceraldehyde-3-phosphate dehydrogenase gene; thrombosis prevention;	
P	(CENG) CENTRAL GLASS CO LTD.		KW	thrombosis treatment.	
P	(TOYO) TOYO SODA MFG CO LTD.		XX	Homo sapiens.	
P	(NISC) NISSAN CHEM IND LTD.		OS		
P	(NISC) NISSAN CHEMICAL INDS KK.		XX		
X	Tagawa M, Wada M, Yamada M, Yokoyama M, Numao N;		XX		
X	WPI; 1987-222882/32.		Key	Location/Qualifiers	
D	N-PSDB; AAN70390.		FT	Peptide	1. .20
D			FT	Protein	21. .411
X	Hybrid plasminogen activator-like polypeptide - having a region for		XX		
F	affinity to fibrin from tissue plasminogen activator and a region from		XX		
P	prourokinase.		PN	EP288435-A.	
X	Disclosure; Fig 2(1-5); 64pp; English.		XX		
P	The TPA portion of the claimed hybrid polypeptide (see FT) may consist of		PD	26-OCT-1988.	
C	2 kringles from N-terminal first serine to 219th glycine of human TPA, 1		XX		
C	kringle from 128th serine to 219th glycine of human TPA or half a kring		PR	15-APR-1987; 87GB-00009081.	
C	le from 161st methionine to 219th glycine (see AAP70257). The C-terminal		PR	16-JUN-1987; 87GB-00014059.	
C	half of the hybrid polypeptide may contain an AA sequence from 150th		PR	04-DEC-1987; 87IE-00003239.	
C	glutamine to C-terminal 411th leucine of prourokinase (see AAP70258).		XX	(CIBA) CIBA GEIGY AG.	
C	(Updated on 25-MAR-2003 to correct PA field.)		PI	Meyhack B, Heim J, Burgi R;	
C	Sequence 431 AA;		XX		
X	Query Match 99.4%; Score 2243; DB 1; Length 431;		DR	WPI; 1988-301440/43.	
S	Best Local Similarity 98.1%; Pred. No. 1.2e-174;		DR	N-PSDB; AAN80981.	
	atches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;		XX		
Q	1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFRG 60		XX		
D	21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFRG 80		XX		
Q	61 KASDTWGRPCLPNSATVLTQTVHAHRSALQLGLGKNYCRNPNRRPPWCYVQVGLK 120		XX		
D	81 KASDTWGRPCLPNSATVLTQTVHAHRSALQLGLGKNYCRNPNRRPPWCYVQVGLK 140		XX		
Q	121 PLVQECMVHDCADGK-----LKFCQGKTLRPRFKIIGGFTTIENQFWFAAIYRRH 172		XX		
D	141 PLVQECMVHDCADGKSPPEELKFCQGKTLRPRFKIIGGFTTIENQFWFAAIYRRH 200		XX		
Q	173 RGSVTVVCGSLISPCWVISATHCFIDYPKEDYIVLGRSLNSNTQEMKFEVENLI 232		XX		
D	201 RGSVTVVCGSLISPCWVISATHCFIDYPKEDYIVLGRSLNSNTQEMKFEVENLI 260		XX		
Q	233 LHKDYSADTLAHHNDIALKIRSEGRCAQPSRTIQTICLPSMYNDPQFTSCITGFGK 292		XX		
D	261 LHKDYSADTLAHHNDIALKIRSEGRCAQPSRTIQTICLPSMYNDPQFTSCITGFGK 320		XX		
Q	293 ENSTDYLPQELKMTVVKLSHRECOQPHYGVSEVTKMLCAADPQWKTSCQGDGGPL 352		XX		
D	321 ENSTDYLPQELKMTVVKLSHRECOQPHYGVSEVTKMLCAADPQWKTSCQGDGGPL 380		XX		
Q	353 VCSLQGRMTLTGIVSWRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403		XX		
D	381 VCSLQGRMTLTGIVSWRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431		XX		
F	JLT 10		XX		
A	30430		XX		
I	AAP80430 standard; protein; 431 AA.		XX		

Query Match 99.4%; Score 2243; DB 1; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFRG 80

QY 61 KASTDTMGRPCLPWNSATVLTQYTHAHRSDALQLGLGKHNKYNCRNPNRRRPPKCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLTQYTHAHRSDALQLGLGKHNKYNCRNPNRRRPPKCYVQVGLK 140
QY 121 PLVQECMWHDCADGK-----LKFCQCGOKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 172
DB 141 PLVQECMWHDCADGKFPSPPELKFQCGOKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 200
QY 173 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGENKFEVENLI 232
DB 201 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGENKFEVENLI 260
QY 233 LHKDYSADTLAHDNDIALKIRSKGRCQAQPSRTIQTICLPMSYNDPQFGTSCITGFGK 292
DB 261 LHKDYSADTLAHDNDIALKIRSKGRCQAQPSRTIQTICLPMSYNDPQFGTSCITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGGFL 352
DB 321 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGGFL 380
QY 353 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RE LT 11
AA 1204
ID AAP81204 standard; protein; 431 AA.
XX
AC AAP81204;
XX
XX 25-MAR-2003 (revised)
DT 03-DEC-1990 (first entry)
XX
XX Pro-urokinase with signal sequence.
XX pro-urokinase (pro-UK); plasminogen activator; PUK33; ss.
XX Homo sapiens.
XX
XX Key : Location/Qualifiers
FH Peptide : 1..20
FT /label= signal peptide
FT Protein : 21..431
FT /label= pro-urokinase
XX
XX EP265874-A.
XX
XX 04-MAY-1988.
XX
XX 23-OCT-1987; 87EP-00115600.
XX
XX 23-OCT-1986; 86JP-00253078.
XX
XX (GREC) GREEN' CROSS CORP.
XX
XX Amatsuji Y, Okabayashi K, Nagai M, Arimura H, Suyama T;
XX
XX WPI: 1988-121000/18.
XX N-PSDB; AN81558.
XX
XX Glycosylated single-chain pro-urokinase prodn. - by cultivating DHFR gene
XX -deficient CHO-K1 cells transformed with a plasmid contg. CDNA, SV40
XX promoter and DHFR gene.
XX
XX Disclosure; Page ?; 19pp; English.
XX
XX The Arg at position 2 is encoded by TGA(sic). Possible error in the
XX specification. Should read CGA ? The pro-UK gene was derived from plasmid
XX PUK33. The cDNA was synthesised using urokinase mRNA isolated from a
XX human kidney cell line. Pro-UK was cloned into a SV40 promoter-contg.
XX plasmid ,down- stream of the promoter. This plasmid was then ligated to

CC a DHFR- gene contg. plasmid so that pro-UK and DHFR are inserted in
CC opposite directions. The recombinant plasmid was used to transform CHO-K1
CC cell derived DHFR gene-deficient host cells to produce glycosylated
CC single-chain pro-UK. (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 1; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNEHQVPSNCDCLNGCTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
DB 21 SNEHQVPSNCDCLNGCTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 80
QY 61 KASTDTMGRPCLPWNSATVLTQYTHAHRSDALQLGLGKHNKYNCRNPNRRRPPKCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLTQYTHAHRSDALQLGLGKHNKYNCRNPNRRRPPKCYVQVGLK 140
QY 121 PLVQECMWHDCADGK-----LKFCQCGOKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 172
DB 141 PLVQECMWHDCADGKFPSPPELKFQCGOKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 200
QY 173 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGENKFEVENLI 232
DB 201 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGENKFEVENLI 260
QY 233 LHKDYSADTLAHDNDIALKIRSKGRCQAQPSRTIQTICLPMSYNDPQFGTSCITGFGK 292
DB 261 LHKDYSADTLAHDNDIALKIRSKGRCQAQPSRTIQTICLPMSYNDPQFGTSCITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGGFL 352
DB 321 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGGFL 380
QY 353 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431
RESULT 12
AAP92119
ID AAP92119 standard; protein; 431 AA.
XX
XX AAP92119;
XX
XX 25-MAR-2003 (revised)
DT 29-JUN-1990 (first entry)
XX
XX Natural human prourokinase.
XX
XX Human prourokinase; antithrombotic; derivative.
XX
XX Homo sapiens.
XX
XX Key : Location/Qualifiers
FH Misc-difference 1
FT /note= "Optional in new deriv."
FT Misc-difference 2..155
FT /note= "Incorporated into new deriv."
FT Misc-difference 135
FT /note= "May be replaced by a non-basic AA in new deriv."
FT Misc-difference 156
FT /note= "Undefined residue in new deriv."
FT Misc-difference 157
FT /note= "Pro, Gly, Ala or Val in new deriv."
FT Misc-difference 158
FT /note= "Lys or Arg in new deriv."
XX
XX W08901513-A.
XX
XX 23-FEB-1989.
XX
XX

18-AUG-1988; 88WO-JP000815.
 19-AUG-1987; 87JP-00204149.
 (SAGA) SAGAMI CHEM RES CENTRE.
 (CENG) CENTRAL GLASS CO LTD.
 (HODO) HODOGAYA CHEM KK.
 (NIPS) NIPPON SODA CO.
 (NISC) NISSAN CHEM IND LTD.
 Kobayashi Y., Omori M., Yamada C;
 WPI; 1989-068869/09.
 N-PSDB; AAN91075.
 Antithrombotic fast-acting pro-urokinase deriv. - produced by culture of
 E coli transformant contg. new plasmid of PMUT9Q family.
 Disclosure; Fig 1; 75pp; Japanese.
 A human prourokinase (PU) deriv. is new which is based upon residues 2-
 155 of natural human prourokinase. The new deriv. is produced by E. coli
 J103/PMUT9Q-RPX in culture. It is a fast-acting drug for the treatment
 and prevention of thrombosis. (Updated on 25-MAR-2003 to correct PA
 field.) (Updated on 25-MAR-2003 to correct PI field.)
 Sequence 431 AA;
 Query Match 99.4%; Score 2243; DB 1; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 atches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 1 SNELHVPNSCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 60
 21 SNELHVPNSCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 80
 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHYCRNPNRRRPPWCYVQVGLK 120
 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHYCRNPNRRRPPWCYVQVGLK 140
 121 PLVQECMVHDCADGK-----LKFCQCKTLRPFKIIIGGEFTTIENQPFALYRRH 172
 141 PLVQECMVHDCADGKPPSPPEELKFCQCKTLRPFKIIIGGEFTTIENQPFALYRRH 200
 173 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 232
 201 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 260
 233 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCBITGFGK 292
 261 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCBITGFGK 320
 293 ENSTDYLYPEQLKMTVVVKLI SHRECCQPHYGSEVTTKMLCAADPOWKTDSCQDGGPL 352
 321 ENSTDYLYPEQLKMTVVVKLI SHRECCQPHYGSEVTTKMLCAADPOWKTDSCQDGGPL 380
 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431
 13
 07112
 AAR07112 standard; protein; 431 AA.
 AAR07112;
 25-MAR-2003 (revised)
 24-JAN-1991 (first entry)
 Human pro-Urokinase encoded by plasmid pUK1.
 pro-Urokinase; transgenic mice.

XX OS Synthetic.
 XX PN EP390592-A.
 XX PD 03-OCT-1990.
 XX PF 30-MAR-1990; 90EP-00303445.
 XX PR 31-MAR-1989; 89JP-00078574.
 XX PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX PA (EXPE-) CENT INST EXPR ANIMALS.
 XX PA (JIKK-) JIKKEN DOBUTSU CHOO KENK.
 XX PI Sekine S, Ito S, Katsuki M;
 XX DR WPI; 1990-299492/40.
 XX DR N-PSDB; AAO06049.
 XX PT Prodn. of recombinant protein, esp. human pro:urokinase - from milk of
 XX PT transgenic animals using promoter of bovine alpha S1 casein chromosomal
 XX PS gene.
 XX PS Example; Table 1; 55pp; English.
 XX CC E. coli strain C600SF8 was transformed with recombinant plasmid containing
 XX CC ds DNA derived from human pharynx cancer cell strain Detroit 562. 10000
 XX CC colonies were screened and one positive clone was identified. Plasmid
 XX CC pUK1 was isolated and found to contain the coding region and 3' non-
 XX CC coding region of pro-UK downstream of Cys(41). Four silent substitutions
 XX CC were identified c.f. Holmes et al., Biotechnology, vol.3, p.923 (1985) as
 XX CC follows: (sic) Asn(254), AAC to AAT; Leu(340), CTA to CTG; Pro(345), CCC
 XX CC to CCA; Gln(346), CAA to CAG. See also AAO06045-Q06048 and AAO06332.
 XX CC (Updated on 25-MAR-2003 to correct PA field.)
 XX SQ Sequence 431 AA;
 Query Match 99.4%; Score 2243; DB 2; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNELHVPNSCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 60
 DB 21 SNELHVPNSCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 80
 QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHYCRNPNRRRPPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHYCRNPNRRRPPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKFCQCKTLRPFKIIIGGEFTTIENQPFALYRRH 172
 DB 141 PLVQECMVHDCADGKPPSPPEELKFCQCKTLRPFKIIIGGEFTTIENQPFALYRRH 200
 QY 173 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 232
 DB 201 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCBITGFGK 292
 DB 261 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCBITGFGK 320
 QY 293 ENSTDYLYPEQLKMTVVVKLI SHRECCQPHYGSEVTTKMLCAADPOWKTDSCQDGGPL 352
 DB 321 ENSTDYLYPEQLKMTVVVKLI SHRECCQPHYGSEVTTKMLCAADPOWKTDSCQDGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
 DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431
 RESULT 14
 AAR04253

AA04253 standard; protein; 431 AA.
AA04253;
25-MAR-2003 (revised)
12-SEP-1990 (first entry)
Human pro-urokinase from the cDNA of clone pcUK176.
Non-glycosylated, pro-urokinase; E. coli; Ptp promoter; MS-2 RBS.
Synthetic.
EP365894-A.
02-MAY-1990.
06-OCT-1989; 89EP-00118586.
11-OCT-1988; 88GB-00023833.
(FARM) FARMITALIA ERBA SPA CARLO.
Brandazza A, Sarmientos P, Orsini G;
WPI; 1990-133447/18.
N-PSDB; AAQ04107.
Non-glycosylated pro-urokinase prodn. - using E coli B strains and E coli promoter PTPR and Shine-Dalgarno sequence MS-2.
Disclosure; Page ?; -pp; English.
SER residue at position 21 is the start of the mature proUK. Non-glycosylated proUK (MW 45KD) produced by E.coli B strain containing the sequence. . See also AAQ04101-07. (Updated on 25-MAR-2003 to correct PA field.)
Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 2; Length 431;
st Local Similarity 98.1%; Pred. No. 1.2e-174;
ches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
1 SNEHQVPSNCDCLNGTCVSNKYFNSHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
21 SNEHQVPSNCDCLNGTCVSNKYFNSHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
61 KASDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNCRPNRRPWCYVQVGLK 120
81 KASDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNCRPNRRPWCYVQVGLK 140
121 PLVQECMVHDCADGK-----LKFOGQKTLRPREKIIGBFTTIENQFWAAYRRH 172
141 PLVQECMVHDCADGKPPSPPEELKFCQGQKTLRPREKIIGBFTTIENQFWAAYRRH 200
173 RGGSVYVCGGSLSPCWVISATHCFIDYPRKEDIYVILGRSLNSNTQGMKFEVENLI 232
201 RGGSVYVCGGSLSPCWVISATHCFIDYPRKEDIYVILGRSLNSNTQGMKFEVENLI 260
233 LHKDYSADTLAHHNDIALLKIRSEKCAQPSRTIQICLPWYNDPQFTSCETGFGK 292
261 LHKDYSADTLAHHNDIALLKIRSEKCAQPSRTIQICLPWYNDPQFTSCETGFGK 320
293 ENSDYLYPEQLKMTVVKVLIHRECCQPHYVGSVTTKMLCAADPQWKTDCSQGDSGGL 352
321 ENSDYLYPEQLKMTVVKVLIHRECCQPHYVGSVTTKMLCAADPQWKTDCSQGDSGGL 380
353 VCSLQGRMTLTGIVSWGRCALDKDPGYTVRVSHFLPWIRSHYKENGALAL 403
381 VCSLQGRMTLTGIVSWGRCALDKDPGYTVRVSHFLPWIRSHYKENGALAL 431

RESULT 15
AA063141
ID AA063141 standard; protein; 431 AA.
XX
AC AA063141;
XX
DT 25-MAR-2003 (revised)
DT 09-JUN-1995 (first entry)
XX
DE Full length human urokinase protein.
XX
KW Human urokinase glycoproteins; cardiovascular diseases;
KW pulmonary embolism.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Sig_peptide 1..20
FT Disulfide-bond 70..151
FT Disulfide-bond 91..133
FT Disulfide-bond 122..146
FT Disulfide-bond 168..299
FT Cleavage-site 179..180
FT /note= "cleavage of this site produces a bioactive two
FT chain form of urokinase"
FT Disulfide-bond 209..225
FT Disulfide-bond 217..288
FT Disulfide-bond 313..382
FT Disulfide-bond 345..361
FT Disulfide-bond 372..400
XX
FN EP620279-A1.
XX
XX 19-OCT-1994.
XX
PF 14-APR-1983; 94EP-00104777.
XX
PR 15-APR-1983; 82US-00368773.
PR 14-MAR-1983; 83US-00474930.
PR 14-APR-1983; 83EP-00103629.
XX
XX (GETH) GENENTECH INC.
XX
PI Heyneker HL, Holmes WE, Vehar GA;
XX
XX WPI: 1994-318362/40.
DR N-PSDB; AAQ73483.
XX
XX Prodn. of human urokinase glycoproteins - using a recombinant expression
PT system used for the treatment of vascular diseases or conditions.
XX
PS Claim 1; Fig 4; 4lpp; English.
XX
CC AAQ73483 is the cDNA sequence which encodes AA063141 the full length
CC 54000 dalton human urokinase (UK) protein. This cDNA was used in the
CC construction of a plasmid capable of transforming either yeast or
CC vertebrate cells, enabling them to produce the 54000 dalton human UK
CC protein. The UK glycoprotein produced could then be used in the treatment
CC of cardiovascular diseases, including pulmonary embolism. The UK produced
CC using this method had the advantage of a specific activity towards fibrin
CC and extant thrombi, not demonstrated previously with UK isolated from
CC natural sources. (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct
CC PR field.)
XX
SQ Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 2; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNEHQVPSNCDCLNGTCVSNKYFNSHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
|||||

21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIEDKSKTCYEGNGHYRG 80
 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKKNYCRPNRRPWCYVQVGLK 120
 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKKNYCRPNRRPWCYVQVGLK 140
 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQPFPAALYYRH 172
 141 PLVQECMVHDCADGKSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFPAALYYRH 200
 173 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 232
 201 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 260
 233 LHKDYSADTLAHHNDIALALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGK 292
 261 LHKDYSADTLAHHNDIALALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGK 320
 293 ENSTDYLYPEQLKMTVVVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDSDGGPL 352
 321 ENSTDYLYPEQLKMTVVVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDSDGGPL 380
 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 16
 AAY50869
 ID AAY50869 standard; protein; 431 AA.
 XX
 AC AAY50869;
 XX
 DT 13-JUL-1994 (first entry)
 XX
 DE Pro-urokinase derivative.
 XX
 KW Pro-urokinase; half-life; thrombolytic; thrombosis; fibrinolytic; factor.
 XX
 OS Homo sapiens.
 XX
 PN JP05336965-A.
 XX
 PD 21-DEC-1993.
 XX
 PF 17-OCT-1991; 91JP-00269615.
 XX
 PR 17-OCT-1991; 91JP-00269615.
 XX
 PS (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI WPI; 1994-030907/04.
 XX
 DR N-PSDB; AAQ55772.
 XX

Novel human pro-urokinase derivs. having long half-life - with high
 thrombolytic activity, useful for treatment of thrombosis.
 XX
 PS Disclosure; Page 15-17; 29pp; Japanese.
 XX

Sequences (AAQ55771-72) are pro-urokinase derivatives. The products have
 an inserted sugar moiety having an amino acid substituted, depleted or
 inserted variant around the thrombin cleavage site. They also have a long
 half-life allowing them to be used in the treatment of thrombosis
 XX
 PS Sequence 431 AA;
 XX

Query Match 99.4%; Score 2243; DB 2; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIEDKSKTCYEGNGHYRG 80

61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKKNYCRPNRRPWCYVQVGLK 120
 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKKNYCRPNRRPWCYVQVGLK 140
 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQPFPAALYYRH 172
 141 PLVQECMVHDCADGKSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFPAALYYRH 200
 173 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 232
 201 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 260
 233 LHKDYSADTLAHHNDIALALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGK 292
 261 LHKDYSADTLAHHNDIALALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGK 320
 293 ENSTDYLYPEQLKMTVVVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDSDGGPL 352
 321 ENSTDYLYPEQLKMTVVVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDSDGGPL 380
 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 17
 AAY50869
 ID AAY50869 standard; protein; 431 AA.
 XX
 AC AAY50869;
 XX
 DT 24-FEB-2000 (first entry)
 XX
 DE Human urokinase protein fragment.
 XX
 KW Urokinase; human; thrombolytic agent; streptokinase; antigenic;
 XX
 KW blood clot; heart attack; treatment.
 XX
 OS Homo sapiens.
 XX
 PN WO9957251-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 06-MAY-1999; 99WO-US010086.
 XX
 PR 06-MAY-1998; 98US-0084392P.
 XX
 PS (OKLA-) OKLAHOMA MEDICAL RES FOUND.
 XX
 PI Zhang XC, Lin X, Tang JYN;
 XX
 DR WPI; 2000-052966/04.
 XX

New thrombolytic agents derived from modified humanized streptokinase,
 useful for treating blood clot disorders.
 XX
 PS Disclosure; Page 46-48; 55pp; English.
 XX

This invention describes a novel thrombolytic agent comprising
 streptokinase where at least one nonessential portion has been modified.
 The invention also describes a method of forming a thrombolytic agent
 comprising determining a nonessential portion of streptokinase and
 modifying the nonessential portion to render the resulting protein less
 antigenic. The modified streptokinase is used to treat blood clot
 disorders, such as heart attacks. The modified streptokinase has less
 antigenicity than streptokinase but is still able to complex plasminogen
 and lead to plasminogen activation. Modified streptokinase with the
 nonessential portions removed or truncated simplify the molecule. Such
 smaller proteins are cheaper and easier to produce. This sequence
 represents a fragment of the human urokinase protein which is used in the
 description of the method of the invention

Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 3; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 80
61 KASTDTMGPCPLPWSNATVLOQTYHAHRSALQGLGKHNYCRNPNRRRRCVYQVGLK 120
81 KASTDTMGPCPLPWSNATVLOQTYHAHRSALQGLGKHNYCRNPNRRRRCVYQVGLK 140
121 PLVQECWVHDCADGK-----LKFCQCKTLRPFKIIIGGEFTTIENQWFAAIYRRH 172
141 PLVQECWVHDCADGKPPSPPELAFQCGQKTLRPFKIIIGGEFTTIENQWFAAIYRRH 200
173 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFVENLI 232
201 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFVENLI 260
233 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCITGFGK 292
261 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCITGFGK 320
293 ENSTDYLYPEQLKMTVVKLIHSHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 352
321 ENSTDYLYPEQLKMTVVKLIHSHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 380
353 VCSLQGRWTLTGIVSWGRCALKDHPGVYTRVSHFLPWIRSHTKKEENGLAL 403
381 VCSLQGRWTLTGIVSWGRCALKDHPGVYTRVSHFLPWIRSHTKKEENGLAL 431
RE LT 18
AF 9591
ID AAY99591 standard; protein; 431 AA.
AC AAY99591;
XX 13-SEP-2000 (first entry)
DT Human plasminogen activator urokinase, u-PA.
DE Human; serine protease; plasminogen activator; cardiant; thrombolytic;
XX heart attack; stroke; blood clotting disorder.
XX Homo sapiens.
XX WO200032759-A1.
XX 08-JUN-2000.
XX 06-MAY-1999; 99WO-US009991.
XX 02-DEC-1998; 98US-0110588P.
XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX Lin X, Zhang XC, Tang JUN;
XX WPI; 2000-422975/36.
XX Polypeptide with plasminogen activator activity useful as thrombolytic
XX agent for treating blood clot disorders e.g. heart attack, comprises 10
XX amino acid peptide fragment for recognition or activation of plasminogen.
XX Disclosure; Page 26-28; 41pp; English.
XX The present sequence is human plasminogen activator urokinase (u-PA), a
XX serine protease which hydrolyses a peptide bond in human plasminogen to

convert it to its active form, plasmin. plasminogen is the principal
serine protease zymogen in the extracellular fluids of vertebrates.
plasmin is implicated in pericellular proteolysis associated with a wide
range of physiological and pathological processes. Plasminogen activators
regulate plasminogen expression either by hydrolysing a peptide bond, as
in the case of u-PA, or by forming tight binding complexes with
plasminogen to spontaneously convert it to plasmin. Review of sequence
homologies of several plasminogen activators and chymotrypsin has
identified a six amino acid peptide involved in plasminogen activation.
This peptide is particularly useful when inserted between amino acid
residues 64 and 84 of full length human plasminogen. Novel plasminogen
activators have been made based upon the plasminogen
activation/recognition site of plasminogen binding proteins. The
polypeptides are useful in preparing thrombolytic agents for treating
blood clotting disorders such as heart attack
Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 3; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 80
61 KASTDTMGPCPLPWSNATVLOQTYHAHRSALQGLGKHNYCRNPNRRRRCVYQVGLK 120
81 KASTDTMGPCPLPWSNATVLOQTYHAHRSALQGLGKHNYCRNPNRRRRCVYQVGLK 140
121 PLVQECWVHDCADGK-----LKFCQCKTLRPFKIIIGGEFTTIENQWFAAIYRRH 172
141 PLVQECWVHDCADGKPPSPPELAFQCGQKTLRPFKIIIGGEFTTIENQWFAAIYRRH 200
173 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFVENLI 232
201 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFVENLI 260
233 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCITGFGK 292
261 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCITGFGK 320
293 ENSTDYLYPEQLKMTVVKLIHSHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 352
321 ENSTDYLYPEQLKMTVVKLIHSHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 380
353 VCSLQGRWTLTGIVSWGRCALKDHPGVYTRVSHFLPWIRSHTKKEENGLAL 403
381 VCSLQGRWTLTGIVSWGRCALKDHPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 19
AAB84605
ID AAB84605 standard; protein; 431 AA.
XX AAB84605;
XX 05-SEP-2001 (first entry)
DT Amino acid sequence of urokinase plasminogen activator.
DE Growth factor; protein inhibitor; protease; damaged tissue;
XX platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;
XX connective tissue derived growth factor; CTGF; chrysalin; VEGF;
XX keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;
XX transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;
XX granulocyte macrophage colony stimulating factor; GM-CSF; uPA;
XX vascular endothelial growth factor; urokinase plasminogen activator;
XX dermal ulcer; wound.
XX Homo sapiens.
XX WO200149309-A2.

X F 12-JUL-2001.
X A 21-DEC-2000; 2000WO-IB001935.
X Y 29-DEC-1999; 99GE-00030768.
X F (PFI) PFIZER LTD.
X F (PFI) PFIZER INC.
X F Davies MJ, Huggins JP, McIntosh FS, Occleston NL;
X WPI; 2001-418351/44.
X L N-PSDB; AAH28220.
X F Composition for the treatment of damaged tissue i.e. chronic wounds and
X F dermal ulcers comprises an inhibitor agent i.e. a protease and a growth
X F factor.
X F Disclosure; Page 550; 572pp; English.
X C The specification describes a pharmaceutical composition, comprising a
X C growth factor, an inhibitor agent, i.e. a protease. The inhibitor agent
X C inhibits the action of at least one specific adverse protein, i.e. a
X C protease, that is upregulated in a damaged tissue such as a wound
X C environment. Growth factors which are included in the composition of the
X C invention are platelet-derived growth factor (PDGF), fibroblast growth
X C factor (FGF), connective tissue derived growth factor (CTGF),
X C keratinocyte-derived growth factor (KGF), transforming growth factor-beta
X C (TGF-beta), granulocyte macrophage colony stimulating factor (GM-CSF),
X C epidermal growth factor (EGF), vascular endothelial growth factor (VEGF),
X C and chrysalin. Inhibitors which are included in the composition of the
X C invention include inhibitors of urokinase-type plasminogen activator
X C (uPA) and matrix metalloproteinase (MMP). The composition is useful for
X C the treatment of chronic damaged tissue, i.e. wounds and dermal ulcers.
X C The present sequence represents a human uPA, and is used to produce the
X C composition of the invention
X S Sequence 431 AA;
X S Query Match 99.4%; Score 2243; DB 4; Length 431;
X S Best Local Similarity 98.1%; Pred. No. 1.2e-174;
X S Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
X C 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
X F 21 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 80
X C 61 KASDTMTGRPCLPWNSATVLTQTYHAHRSALQGLGKHNCRPNRRRPPWCYVQVGLK 120
X L 81 KASDTMTGRPCLPWNSATVLTQTYHAHRSALQGLGKHNCRPNRRRPPWCYVQVGLK 140
X C 121 PLVQECMWHDCADGK-----LKFCGQKTLRPRFKIIGGFTTIENQWPAAIYRRH 172
X F 141 PLVQECMWHDCADGKSPPEELKFCGQKTLRPRFKIIGGFTTIENQWPAAIYRRH 200
X C 173 RGGSVTVVCGSLISPCWVISAHCIFDYPKEDYIVLGRSLNNTQGMKFEVENLI 232
X D 201 RGGSVTVVCGSLISPCWVISAHCIFDYPKEDYIVLGRSLNNTQGMKFEVENLI 260
X C 233 LHKDYSADTLAHNDIALKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
X F 261 LHKDYSADTLAHNDIALKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
X C 293 ENSTDYLYPQLKMTVVKLISHRECQPHYVSEVTMCLCAADPQWKTDSOCGDSGGL 352
X L 321 ENSTDYLYPQLKMTVVKLISHRECQPHYVSEVTMCLCAADPQWKTDSOCGDSGGL 380
X C 353 VCSLQGRMTLTGIVSWRGKCALDKPGVYTRVSHFLPWIRSHTKENGLAL 403
X D 381 VCSLQGRMTLTGIVSWRGKCALDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 20
AAE17128
ID AAE17128 standard; protein; 431 AA.
XX AC AAE17128;
XX DT 18-APR-2002 (first entry)
XX DE Human uPA protein.
XX KW Human; cancer; urokinase-type plasminogen activator; uPA; inflammation;
KW Ets-1 transcription factor; N-acetylglucosaminyltransferase V; Gnt-V;
KW matrix-type metalloproteinase; MMP-1; MMP-3; gene therapy.
XX OS Homo sapiens.
XX PN WO200196606-A2.
XX XX
XX PD 20-DEC-2001.
XX PF 14-JUN-2001; 2001WO-US019248.
XX PR 14-JUN-2000; 2000US-00593488.
XX PA (NYXI-) NYXIS NEURO THERAPIES INC.
XX PI Yamamoto H, Kroes R, Moskal JR;
XX WPI; 2002-130746/17.
XX N-PSDB; AAD27855.
XX Identifying a compound for treating cancer, comprises detecting
XX transcription factor Ets-1, N-acetylglucosaminyltransferase V, urokinase-
XX type plasminogen activator, matrix-type metalloproteinase-1 and -3 gene
XX expression.
XX Example 1; Page 62-63; 63pp; English.
XX The invention relates to a method of identifying a compound for treating
XX cancer. The method involves detecting the expression of a panel of
XX sequences selected from transcription factor Ets-1, urokinase-type
XX plasminogen activator (uPA), N-acetylglucosaminyltransferase V (Gnt-V),
XX matrix-type metalloproteinase (MMP)-1 and MMP-3 in the cell. The method
XX is useful for identifying a compound that affects a cell, particularly a
XX cancer cell or glioma cell, or a cell that is involved in inflammation.
XX It is used for diagnosing and/or treating cancer or other conditions that
XX are affected by one or more members of a panel of genes or their protein
XX product. The method is also useful for drug discovery, drug safety
XX evaluations and in gene therapy. The present sequence is human uPA
XX protein
XX Sequence 431 AA;
XX Query Match 99.4%; Score 2243; DB 5; Length 431;
XX Best Local Similarity 98.1%; Pred. No. 1.2e-174;
XX Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
X C 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
X D 21 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 80
X C 61 KASDTMTGRPCLPWNSATVLTQTYHAHRSALQGLGKHNCRPNRRRPPWCYVQVGLK 120
X D 81 KASDTMTGRPCLPWNSATVLTQTYHAHRSALQGLGKHNCRPNRRRPPWCYVQVGLK 140
X C 121 PLVQECMWHDCADGK-----LKFCGQKTLRPRFKIIGGFTTIENQWPAAIYRRH 172
X D 141 PLVQECMWHDCADGKSPPEELKFCGQKTLRPRFKIIGGFTTIENQWPAAIYRRH 200
X C 173 RGGSVTVVCGSLISPCWVISAHCIFDYPKEDYIVLGRSLNNTQGMKFEVENLI 232
X D 201 RGGSVTVVCGSLISPCWVISAHCIFDYPKEDYIVLGRSLNNTQGMKFEVENLI 260

Qy 1 SNELHVPNSDCLNGTGVSNKYPFNSNTHWNCPPKFGQHCIEDKSKTCYEGNGHYRG 60
 Di 21 SNELHVPNSDCLNGTGVSNKYPFNSNTHWNCPPKFGQHCIEDKSKTCYEGNGHYRG 80
 Qy 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
 Di 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 140
 Qy 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPRFKIIGGFTTINQWPAAYRRH 172
 Di 141 PLVQECMVHDCADGKPPSPPEELKFCQCGQKTLRPRFKIIGGFTTINQWPAAYRRH 200
 Qy 173 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRNSNTQGMKFEVENLI 232
 Di 201 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRNSNTQGMKFEVENLI 260
 Qy 233 LHKDYSADTLAHHNDIALKIRSKGRCAPQSRRTIQTICLPSMYNDPQGTSCITGFGK 292
 Di 261 LHKDYSADTLAHHNDIALKIRSKGRCAPQSRRTIQTICLPSMYNDPQGTSCITGFGK 320
 Qy 293 ENSTDYLYPEQLKMTVVKLI SHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPL 352
 Di 321 ENSTDYLYPEQLKMTVVKLI SHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPL 380
 Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKXENGLAL 403
 Di 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKXENGLAL 431

RE LT 24
 AF 6547
 II ABUS6547 standard; protein; 431 AA.
 XY ABUS6547;
 AC
 Xy 02-APR-2003 (first entry)
 D3 Lung cancer-associated polypeptide #140.
 D5 Lung cancer-associated polypeptide #140.
 Xy Lung cancer-associated polypeptide; cytostatic; emphysema;
 Xy antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 Xy small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 Xy chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 Xy interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 Xy Unidentified.
 Xy WO200286443-A2.
 Xy 31-OCT-2002.
 Xy 18-APR-2002; 2002WO-US012476.
 Xy 18-APR-2001; 2001US-0284770P.
 Xy 10-MAY-2001; 2001US-0290492P.
 Xy 09-NOV-2001; 2001US-0339245P.
 Xy 13-NOV-2001; 2001US-0350666P.
 Xy 29-NOV-2001; 2001US-0334370P.
 Xy 12-APR-2002; 2002US-0372246P.
 Xy (EOSB-) EOS BIOTECHNOLOGY INC.
 Xy Aziz N, Murray R;
 Xy WPI; 2003-093161/08.
 DE N-PSDB; ABX76275.
 Xy Detecting a lung cancer-associated transcript in a cell from a patient
 Xy for treating lung cancer, by contacting a biological sample from the
 Xy patient with a polynucleotide that exhibits increased or decreased
 Xy expression in lung cancer.

PS Claim 27; Page 296; 453pp; English.
 XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the
 CC invention
 XX
 SQ Sequence 431 AA;
 Query Match 99.4%; Score 2243; DB 6; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 Qy 1 SNELHVPNSDCLNGTGVSNKYPFNSNTHWNCPPKFGQHCIEDKSKTCYEGNGHYRG 60
 Db 21 SNELHVPNSDCLNGTGVSNKYPFNSNTHWNCPPKFGQHCIEDKSKTCYEGNGHYRG 80
 Qy 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
 Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 140
 Qy 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPRFKIIGGFTTINQWPAAYRRH 172
 Db 141 PLVQECMVHDCADGKPPSPPEELKFCQCGQKTLRPRFKIIGGFTTINQWPAAYRRH 200
 Qy 173 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRNSNTQGMKFEVENLI 232
 Db 201 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRNSNTQGMKFEVENLI 260
 Qy 233 LHKDYSADTLAHHNDIALKIRSKGRCAPQSRRTIQTICLPSMYNDPQGTSCITGFGK 292
 Db 261 LHKDYSADTLAHHNDIALKIRSKGRCAPQSRRTIQTICLPSMYNDPQGTSCITGFGK 320
 Qy 293 ENSTDYLYPEQLKMTVVKLI SHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPL 352
 Db 321 ENSTDYLYPEQLKMTVVKLI SHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPL 380
 Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKXENGLAL 403
 Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKXENGLAL 431
 RESULT 25
 ABUS6708
 ID ABUS6708 standard; protein; 431 AA.
 XX
 AC ABUS6708;
 XX
 DT 02-APR-2003 (first entry)
 XX Lung cancer-associated polypeptide #301.
 XX Lung cancer-associated polypeptide; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX

Unidentified.
 WO200286443-A2.
 31-OCT-2002.
 18-APR-2002; 2002WO-US012476.
 18-APR-2001; 2001US-0284770P.
 10-MAY-2001; 2001US-0290492P.
 09-NOV-2001; 2001US-0339245P.
 13-NOV-2001; 2001US-0350866P.
 29-NOV-2001; 2001US-0334370P.
 12-APR-2002; 2002US-0372246P.
 (BOSB-) BOS BIOTECHNOLOGY INC.
 Aziz N, Murray R;
 WPI; 2003-093161/08.
 N-PSDB; ABX76437.
 Detecting a lung cancer-associated transcript in a cell from a patient for treating lung cancer, by contacting a biological sample from the patient with a polynucleotide that exhibits increased or decreased expression in lung cancer.
 Claim 27; Page 424; 453pp; English.
 The invention relates to a method for detecting a lung cancer-associated transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence that is at least 80 % identical to a gene that exhibits increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the invention.
 Sequence 431; AA;
 Query Match 99.4%; Score 2243; DB 6; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 1 SNELHQPVSNCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFFYRG 60
 21 SNELHQPVSNCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFFYRG 80
 61 KASTDTMGRPCLPWNSATVLOQTHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK 120
 81 KASTDTMGRPCLPWNSATVLOQTHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK 140
 121 PLVQECMVHDCADGK-----LKFCQCKTLRPRFKIIGGFTTIENQWFAAAYRRH 172
 141 PLVQECMVHDCADGKXSSPEELKFCQCKTLRPRFKIIGGFTTIENQWFAAAYRRH 200
 173 RGSSTVYVCGSLSPCWVLSATHCFIDYPKKEDIYVLGRSLNLTQGMKFEVENLI 232
 201 RGSSTVYVCGSLSPCWVLSATHCFIDYPKKEDIYVLGRSLNLTQGMKFEVENLI 260
 233 LHKDYSADTLAHHNDIALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFQTSCEITGFGK 320
 293 ENSTDYLYPEQLKMTVVVKLISHRECOQPHYVYGVSEVTTMLCAADPQWKTDSCQDSSGGL 352
 321 ENSTDYLYPEQLKMTVVVKLISHRECOQPHYVYGVSEVTTMLCAADPQWKTDSCQDSSGGL 380
 353 VCSLQGRMTLTGIVSGRGKCALDKDPGVYTRVSHFLPWIRSHYKEENGLAL 403
 381 VCSLQGRMTLTGIVSGRGKCALDKDPGVYTRVSHFLPWIRSHYKEENGLAL 431
 RESULT 26
 ABUS11076
 ID ABUS11076 standard; protein; 431 AA.
 XX
 AC ABUS11076;
 XX
 DT 05-FEB-2003 (first entry)
 XX
 DE Human urokinase plasminogen activator.
 XX
 KW Urokinase plasminogen activator; gene therapy; cancer;
 KW hyperproliferative disorder; cancer; breast cancer; colon cancer;
 KW bone cancer; brain cancer; ovary cancer; cervix cancer;
 KW endometrium cancer; stomach cancer; kidney cancer; tumour metastasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200279515-A1.
 XX
 PD 10-OCT-2002.
 XX
 PF 18-MAR-2002; 2002WO-US008112.
 XX
 PR 30-MAR-2001; 2001US-00821972.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI Baker BP, Freier SM, Watt AT;
 XX
 WPI; 2003-058441/05.
 DR N-PSDB; ABX17681.
 XX
 PT New antisense compound, useful for preparing a composition for treating hyperproliferative disorders, cancer e.g., breast, colon, bone, brain, ovary, cervix, endometrium, stomach or kidney cancer, or tumor metastasis.
 XX
 PS Disclosure; Page 101-102; 153pp; English.
 XX
 CC A new compound, which is 8-50 nucleobases in length targeted to a nucleic acid molecule encoding urokinase plasminogen activator, specifically CC hybridizes with and inhibits the expression of urokinase plasminogen CC activator. The compound is useful for preparing a composition for CC treating (e.g. by gene therapy) hyperproliferative disorder, cancer e.g., CC breast, colon, bone, brain, ovary, cervix, endometrium, stomach or kidney CC cancer, or tumour metastasis. This is the amino acid sequence of a CC urokinase plasminogen activator
 XX
 SQ Sequence 431 AA;
 Query Match 99.4%; Score 2243; DB 6; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 1 SNELHQPVSNCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFFYRG 60
 21 SNELHQPVSNCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFFYRG 80
 61 KASTDTMGRPCLPWNSATVLOQTHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK 120
 81 KASTDTMGRPCLPWNSATVLOQTHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK 140

QY 121 PLVQECMWHDCADGK-----LKFOGQKTLRPRFKIIGBEFTTIENQFWFAAIYRRH 172
D- 141 PLVQECMWHDCADGKPPSPPEELKFOGQKTLRPRFKIIGBEFTTIENQFWFAAIYRRH 200
QY 173 RGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 232
D- 201 RGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALKIRSEKGRCAQPSRTIOTICLPSMNDPQFSTCEITGFGK 292
D- 261 LHKDYSADTLAHHNDIALKIRSEKGRCAQPSRTIOTICLPSMNDPQFSTCEITGFGK 320
QY 293 ENSTDYLYPEOLKMTVVKVLIHRECCQPHYVYGVSEVTTKMLCAADPOWKTDSCQDGGPL 352
D- 321 ENSTDYLYPEOLKMTVVKVLIHRECCQPHYVYGVSEVTTKMLCAADPOWKTDSCQDGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHYKEENGLAL 403
D- 381 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHYKEENGLAL 431

RE 127
AE 2137
IL ABR92137 standard; protein; 431 AA.
XX ABR92137;
XX 10-SEP-2003 (first entry)
XX Human cervical cancer cell marker protein SEQ ID NO:184.
XX Human; cervical cancer; cervical cancer marker; cancer therapy;
XX detection; gene therapy; vaccine.
XX Homo sapiens.
XX WO2002101075-A2.
XX 19-DEC-2002.
XX 12-JUN-2002; 2002WO-US018638.
XX 13-JUN-2001; 2001US-0298155P.
XX 13-JUN-2001; 2001US-0298159P.
XX 14-NOV-2001; 2001US-033936P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Schlegel R, Chen Y, Zhao X, Monahan JE, Kamatkar S;
P1 Gannavarapu M, Glatt K, Hoersch S;
XX
D- WPI; 2003-156967/15.
D- N-PSDB; ACF12920.
XX
XX New isolated nucleic acid molecule useful for detecting, characterizing,
P1 preventing and treating human cervical cancers, in various prognostic and
P1 diagnostic assays, in pharmacogenomics and in monitoring clinical trials.
XX
XX Claim 4; Page 345-346; 386pp; English.
XX
XX ACF12828 to ACF12947 encode the human cervical cancer marker proteins (I)
CC given in ABR92047 to ABR92164. A higher level of expression of (I) than
CC normal indicates the presence of cervical cancer. Also described: (1) a
CC vector (II) containing (I); (2) a host cell (III) containing (I); and (3)
CC assessing (M1) whether a patient is afflicted with cervical cancer,
CC comprising comparing the level of expression of a marker in a patient's
CC sample, and the normal level of expression of the marker in a control non
CC -cervical cancer sample, where a significant increase in the level of
CC expression of the marker in the patient's sample relative to that in the
CC control sample is an indication that the patient is afflicted with
CC cervical cancer. (I) has cytostatic activity, and can be used in gene
CC therapy and in vaccines. (I) is useful in detecting, characterizing,
CC preventing and treating human cervical cancers. (I) may also be used in

CC various prognostic and diagnostic assays, pharmacogenomics and in
CC monitoring clinical trials
XX
XX Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 6; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPGQHCIEDKSKTCYEGNGHFYRG 60
D- 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
D- 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 140
QY 121 PLVQECMWHDCADGK-----LKFOGQKTLRPRFKIIGBEFTTIENQFWFAAIYRRH 172
D- 141 PLVQECMWHDCADGKPPSPPEELKFOGQKTLRPRFKIIGBEFTTIENQFWFAAIYRRH 200
QY 173 RGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 232
D- 201 RGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALKIRSEKGRCAQPSRTIOTICLPSMNDPQFSTCEITGFGK 292
D- 261 LHKDYSADTLAHHNDIALKIRSEKGRCAQPSRTIOTICLPSMNDPQFSTCEITGFGK 320
QY 293 ENSTDYLYPEOLKMTVVKVLIHRECCQPHYVYGVSEVTTKMLCAADPOWKTDSCQDGGPL 352
D- 321 ENSTDYLYPEOLKMTVVKVLIHRECCQPHYVYGVSEVTTKMLCAADPOWKTDSCQDGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHYKEENGLAL 403
D- 381 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHYKEENGLAL 431

RESULT 28
AAR20537
ID AAR20537 standard; protein; 434 AA.
XX
XX AAR20537;
XX 25-MAR-2003 (revised)
XX 21-MAY-1992 (first entry)
XX
XX Amidated deriv. of pro-urokinase (2).
XX
XX Pro-urokinase; plasminogen activator; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Protein 1. .431
FT /label= pro-UK
XX
XX DE4122688-A.
XX
XX 16-JAN-1992.
XX
XX 09-JUL-1991; 91DE-04122688.
XX
XX 12-JUL-1990; 90GB-00015369.
XX
XX 10-JUL-1991; 91GB-00014846.
XX
XX (FARM) FARMITALIA ERBA SRL CARLO.
XX
XX Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;
P1 WPI; 1992-025815/04.
XX
XX New amidated derivs. of human pro-urokinase - are fibrinolytic and can be

used to treat acute myocardial infarction, pulmonary embolism or deep venous thrombosis.

Claim 4,8; Page 8; 18pp; German.

The protein has fibrinolytic activity and can be used in the same way as PUK, e.g. for treating acute myocardial infarction, lung embolism and deep venous thrombosis. It has greater affinity for plasminogen bound to fibrin than for circulating plasminogen, so have high selectivity for thrombi with reduced chance of bleeding. Compared with the COOH-terminated cpds., it has a better stability against most carboxypeptidases and prolonged half life. See also AAQ20360, AAQ20754-58 and AAR20536-38. (Updated on 25-MAR-2003 to correct PA field.)

Sequence 434 AA;

Query Match 99.4%; Score 2243; DB 2; Length 434;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174; Indels 8; Gaps 1;
 Matches 403; Conservative 0; Mismatches 0;

1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
 61 KASDTMGRPCLPNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRRPPWCYVQVGLK 120
 81 KASDTMGRPCLPNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRRPPWCYVQVGLK 140
 121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGSEFTTIENQPFAAIYRRH 172
 141 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGSEFTTIENQPFAAIYRRH 200
 173 RGSVTVVCGSLSPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 232
 201 RGSVTVVCGSLSPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 260
 233 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
 261 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
 293 ENSTDYLYPEQLKMTVVVKLI SHRECQOPHYVGSSEVTTKMLCAADPQWKTDSCQDSDGGPL 352
 321 ENSTDYLYPEQLKMTVVVKLI SHRECQOPHYVGSSEVTTKMLCAADPQWKTDSCQDSDGGPL 380
 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

ULT 29
 20538
 AAR20538 standard; protein; 434 AA.

AAR20538;
 25-MAR-2003 (revised)
 21-MAY-1992 (first entry)

Amidated deriv. of pro-urokinase (3).

Pro-urokinase; plasminogen activator; ss.

Homo sapiens.

Key Location/Qualifiers
 Protein 1.431
 /label= pro-UK

DE412688-A;
 16-JAN-1992.
 09-JUL-1991; 91DE-04122688.

12-JUL-1990; 90GB-00015369.
 10-JUL-1991; 91GB-00014846.
 (FARM) FARMITALIA ERBA SRL CARLO.

Gozzini L, Visco C, Perago R, Roncucci R, Sarmientos P;
 WPI; 1992-035815/04.

New amidated derivs. of human pro-urokinase - are fibrinolytic and can be used to treat acute myocardial infarction, pulmonary embolism or deep venous thrombosis.

Claim 4,9; Page 8; 18pp; German.

The protein has fibrinolytic activity and can be used in the same way as PUK, e.g. for treating acute myocardial infarction, lung embolism and deep venous thrombosis. It has greater affinity for plasminogen bound to fibrin than for circulating plasminogen, so have high selectivity for thrombi with reduced chance of bleeding. Compared with the COOH-terminated cpds., it has a better stability against most carboxypeptidases and prolonged half life. See also AAQ20360, AAQ20754-58 and AAR20536-38. (Updated on 25-MAR-2003 to correct PA field.)

Sequence 434 AA;

Query Match 99.4%; Score 2243; DB 2; Length 434;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174; Indels 8; Gaps 1;
 Matches 403; Conservative 0; Mismatches 0;

1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
 61 KASDTMGRPCLPNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRRPPWCYVQVGLK 120
 81 KASDTMGRPCLPNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRRPPWCYVQVGLK 140
 121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGSEFTTIENQPFAAIYRRH 172
 141 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGSEFTTIENQPFAAIYRRH 200
 173 RGSVTVVCGSLSPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 232
 201 RGSVTVVCGSLSPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 260
 233 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
 261 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
 293 ENSTDYLYPEQLKMTVVVKLI SHRECQOPHYVGSSEVTTKMLCAADPQWKTDSCQDSDGGPL 352
 321 ENSTDYLYPEQLKMTVVVKLI SHRECQOPHYVGSSEVTTKMLCAADPQWKTDSCQDSDGGPL 380
 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 30
 AAR20536
 ID AAR20536 standard; protein; 436 AA.

AC AAR20536;
 25-MAR-2003 (revised)
 21-MAY-1992 (first entry)

Amidated deriv. of pro-urokinase (1).

Pro-urokinase; plasminogen activator; ss.

Q6 XX Homo sapiens. Location/Qualifiers
Q7 XX Key 1.431
Q8 XX Protein /label= pro-UK
Q9 FT /note= "the amidated form of pro-UK is claimed in claim 1"
Q10 FT Misc-difference 433...436
Q11 FT /note= "may be any amino acid, pref. Lys or Arg, or 0-4 amino acids may be omitted"
Q12 XX DE4122688-A.
Q13 XX 16-JAN-1992.
Q14 XX 09-JUL-1991; 91DE-04122688.
Q15 XX 12-JUL-1990; 90GB-00015369.
Q16 XX 10-JUL-1991; 91GB-00014846.
Q17 XX (FARM) FARMITALIA ERBA SRL CARLO.
Q18 XX Gozzini L, Visco C, Perego R, Roncuetti R, Sarmientos P;
Q19 XX WPI; 1992-025815/04.
Q20 XX New amidated derivs. of human pro-urokinase - are fibrinolytic and can be used to treat acute myocardial infarction, pulmonary embolism or deep venous thrombosis.
Q21 XX Claim 1.4.7; Page 8; 18pp; German.
Q22 XX The protein has fibrinolytic activity and can be used in the same way as PUK, e.g. for treating acute myocardial infarction, lung embolism and deep venous thrombosis. It has greater affinity for plasminogen bound to fibrin than for circulating plasminogen, so have high selectivity for thrombi with reduced chance of bleeding. Compared with the COOH-terminated cpds., it has a better stability against most carboxypeptidases and prolonged half life. See also AAQ20360, AAQ20754-58 and AA20336-38. (Updated on 25-MAR-2003 to correct PA field.)
Q23 XX Sequence 436 AA;
Q24 XX ery Match 99.4%; Score 2243; DB 2; Length 436;
Q25 XX st Local Similarity 98.1%; Pred. No. 1.2e-174;
Q26 XX tches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Q27 Q3 1 SNELHQPNSCDLNGGTCVSNKYFSNTHWCNCPKFGGHCETDKSKTCYEGNGHFPYRG 60
Q28 DI 21 SNELHQPNSCDLNGGTCVSNKYFSNTHWCNCPKFGGHCETDKSKTCYEGNGHFPYRG 80
Q29 Q3 61 KASDTMGRPCLPWNATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRWCYVQVGLK 120
Q30 DI 81 KASDTMGRPCLPWNATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRWCYVQVGLK 140
Q31 Q3 121 PLVQECMVDGADGK-----LKFGCGKTLPRPKIIGGFTTIENOPWFAAIYRRH 172
Q32 DI 141 PLVQECMVDGADGKPPSPPEELKFGCGKTLPRPKIIGGFTTIENOPWFAAIYRRH 200
Q33 Q3 173 RGGSVYVCGSLISPCWVISATCFIDYPKKEDIYVYVGLSRSLNSNTQCEMKEFVENLI 232
Q34 DI 201 RGGSVYVCGSLISPCWVISATCFIDYPKKEDIYVYVGLSRSLNSNTQCEMKEFVENLI 260
Q35 Q3 233 LHKYSADTLAHNDIALKLRSEKRCQAQPSRTIQICLPSMYNDPQFTSCBITGFGK 292
Q36 DI 261 LHKYSADTLAHNDIALKLRSEKRCQAQPSRTIQICLPSMYNDPQFTSCBITGFGK 320
Q37 Q3 293 ENSTDYLYPEQLKMTVVKLI SHRECCQPHYYGSEVTTKMLCAADPQWKTDSCQDSDGGPL 352
Q38 DI 321 ENSTDYLYPEQLKMTVVKLI SHRECCQPHYYGSEVTTKMLCAADPQWKTDSCQDSDGGPL 380
Q39 Q3 353 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 403

DB 381 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 431
RESULT 31
ADE25745
ID ADE25745 standard; protein; 431 AA.
XX AC ADE25745;
XX DT 29-JAN-2004 (first entry)
XX DE Human protein differentially expressed in foam cells #22.
XX KW Human; differential expression; foam cell; LPS; lipopolysaccharide; cardiovascular disease; atherosclerosis.
XX OS Homo sapiens.
XX PN US2003194721-A1.
XX PD 16-OCT-2003.
XX PF 18-SEP-2002; 2002US-00247671.
XX PR 19-SEP-2001; 2001US-0323784P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Mikita T, Shiffman D, Porter JG, Kaser MR;
XX DR WPI; 2003-875398/81.
XX DR N-PSDB; ADE25630.
XX PT Combination containing several polynucleotide that are differentially expressed in foam cells and complements of the polynucleotides, useful for diagnosing cardiovascular disease or atherosclerosis.
XX PS Disclosure; SEQ ID NO 149; 37pp; English.
XX CC The invention relates to a combination comprising several polynucleotides having any one of 127 sequences (S1) such as the sequence of human calmodulin gene, human mRNA for KIA0930 protein, leukotriene A4 hydrolase, human CGI-142 protein mRNA, human K+ channel beat 2 subunit mRNA, etc., and their complements. The cDNAs are differentially expressed in LPS (lipopolysaccharide)-treated foam cells. Also included are obtaining an extended or full length gene from a library of nucleic acid sequences, an expression vector containing the nucleic acids, a host cell containing the vector, a purified polypeptide appearing as ADE25750 and ADE25751, producing a protein by culturing the host cell, and a composition comprising a purified antibody that specifically binds to the proteins. The foam cell-expressed nucleic acids are useful for a high throughput detection of differential expression of one or more polynucleotides in a sample. The sample is from a subject with atherosclerosis and comparison with a standard defines early, mid or late stages of the disorder. The foam cell-expressed nucleic acids are useful for high throughput screening of a library of molecules or compounds to identify a ligand which binds a polynucleotide. The library is chosen from DNA molecules, peptides, proteins and RNA molecules. The protein is useful for a high throughput screening of library of molecules or compounds to identify at least one ligand which specifically binds a protein, for purifying a ligand from a sample for making an antibody. The foam cell-expressed nucleic acids are useful for diagnosing cardiovascular disorder. The foam cell-expressed nucleic acids are useful as elements on a microarray which can be used for detecting related polynucleotide in a sample, diagnosing cardiovascular disease, atherosclerosis. The present sequence represents a protein differentially expressed in LPS treated foam cells.
XX SQ Sequence 431 AA;

Query Match 99.3%; Score 2241; DB 7; Length 431;
Best Local Similarity 97.8%; Pred. No. 1.7e-174;
Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
21 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
61 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKKNYCRNPDRRPPWCYVQVGLK 120
81 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKKNYCRNPDRRPPWCYVQVGLK 140
121 PLVQECMVHDCADGK-----LKFCQGOKTLRPRFKIIGGFTTIENQWFAAIYRRH 172
141 PLVQECMVHDCADGKPPSEELKFCQGOKTLRPRFKIIGGFTTIENQWFAAIYRRH 200
173 RGSVTVVCGSLSPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
201 RGSVTVVCGSLSPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
233 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCIEITGFGK 292
261 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCIEITGFGK 320
293 ENSTDYLYPEQLKMTVVVKLIASHRECQPHYVSGSEVTTKMLCAADPQWKTDSCQDGGPL 352
321 ENSTDYLYPEQLKMTVVVKLIASHRECQPHYVSGSEVTTKMLCAADPQWKTDSCQDGGPL 380
353 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
381 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431
AW24578 standard; protein; 430 AA.
AW24578;
25-MAR-2003 (revised)
11-NOV-1997 (first entry)
Inhibitor resistant urokinase.
Urokinase; inhibitor resistant; plasminogen; human; whey acid protein;
plasminogen activator inhibitor-1; plasminogen activator; protease; WAP;
plasmin; proteolytic enzyme; clot lysis; heart attack; heart muscle.
Homo sapiens;
Synthetic.
Key Location/Qualifiers
Peptide 1..19
Protein /note= "whey acid protein signal peptide"
Domain /note= "urokinase"
Domain 20..64
Domain /note= "B-domain"
Domain 61..150
Domain /note= "Kringlet-1 domain"
Domain 179..430
Domain /note= "P-domain"
Misc-difference 198..203
/note= "deleted in modified urokinase of the invention"
US5648253-A;
15-JUL-1997.
08-SEP-1992; 92US-00942157.
20-DEC-1990; 90US-00631673.
(TSIT-) TSI CORP.
Wei C;

XX WPI; 1997-372062/34.
DR N-PSDB; AAT80075.
XX Deletion-modified urokinase protein - with increased resistance to
PT inhibition by plasminogen activator inhibitor-1.
XX Disclosure; Col 15-18; 16pp; English.
XX This sequence represents the full length urokinase, including the whey
acid protein (WAP) signal peptide. This sequence has residues 179-184 of
the urokinase sequence deleted to create the modified urokinase of the
invention. The modified urokinase (see AAW4578) cleaves plasminogen, and
has a lower binding affinity for plasminogen activator inhibitor-1 than
the corresponding unmodified urokinase. Urokinase is one of two types of
mammalian plasminogen activators (PA), the other being tissue type PA.
PAs catalyze the conversion of the circulating zymogen plasminogen to the
broad spectrum protease plasmin by limited proteolysis. The modified
urokinase can be used for clot lysis, specifically to dissolve heart
attack-causing clots before they cause permanent damage to heart muscle.
The urokinase mutant is more resistant to inhibition by plasminogen
activator inhibitors than the unmodified urokinase. It can be selectively
expressed and secreted from the mammary glands of transgenic animals.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 430 AA;
Query Match 99.2%; Score 2240; DB 2; Length 430;
Best Local Similarity 97.8%; Pred. No. 2.1e-174;
Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 20 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 79
QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKKNYCRNPDRRPPWCYVQVGLK 120
DB 80 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKKNYCRNPDRRPPWCYVQVGLK 139
QY 121 PLVQECMVHDCADGK-----LKFCQGOKTLRPRFKIIGGFTTIENQWFAAIYRRH 172
DB 140 PLVQECMVHDCADGKPPSEELKFCQGOKTLRPRFKIIGGFTTIENQWFAAIYRRH 199
QY 173 RGSVTVVCGSLSPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
DB 200 RGSVTVVCGSLSPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 259
QY 233 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCIEITGFGK 292
DB 260 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCIEITGFGK 319
QY 293 ENSTDYLYPEQLKMTVVVKLIASHRECQPHYVSGSEVTTKMLCAADPQWKTDSCQDGGPL 352
DB 320 ENSTDYLYPEQLKMTVVVKLIASHRECQPHYVSGSEVTTKMLCAADPQWKTDSCQDGGPL 379
QY 353 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
DB 380 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 430
RESULT 33
AAP91886
ID AAP91886 standard; protein; 431 AA.
XX
AC AAP91886;
XX
DT 25-MAR-2003 (revised)
DT 31-OCT-2002 (revised)
DT 16-APR-1990 (first entry)
XX
DE Sequence of prourokinase.
XX
KW Low mol. wt. plasminogen activator; prourokinase; dissolve blood clot.

XX O6 Homo sapiens.
XX O7 Location/Qualifiers
XX O8 21..431
XX O9 /note= "Mature prourokinase."
XX O10 170..179
XX O11 /note= "Preferred initiation region for the low mol. wt.
XX O12 plasminogen activators."
XX O13
XX O14 EP316068-A.
XX O15
XX O16 17-MAY-1989.
XX O17
XX O18 07-OCT-1988; 88EP-00309417.
XX O19
XX O20 09-OCT-1987; 87US-00107370.
XX O21 27-SEP-1988; 88US-00248727.
XX O22
XX O23 (COLB) COLLABORATIVE RES INC.
XX O24
XX O25 Mao JI;
XX O26
XX O27 WPI; 1989-146601/20.
XX O28 N-PSDB; AAN91740.
XX O29
XX O30 Modified low mol. wt. plasminogen activator - formed of amino acids
XX O31 comprising the amino acid portion of prourokinase from 150 to 411.
XX O32
XX O33 Fig 1; Page -; 27pp; English.
XX O34
XX O35 A low mol. wt. polypeptide plasminogen activator (PA) is claimed which is
XX O36 formed from amino acids 150-411 of prourokinase. The preferred initiation
XX O37 region for the low mol. wt. PA is indicated (see FT). A low mol. wt. PA
XX O38 can be injected into blood in the body in vivo to dissolve clots without
XX O39 harm. (Updated on 31-OCT-2002 to add missing OS field.) (Updated on 25-
XX O40 MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI
XX O41 field.)
XX O42
XX O43 Sequence 431 AA;
XX O44
XX O45 Query Match 99.2%; Score 2240; DB 1; Length 431;
XX O46 st Local Similarity 97.8%; Pred. No. 2.1e-174;
XX O47 tches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;
XX O48
XX O49 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHYRG 60
XX O50 21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHYRG 80
XX O51
XX O52 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
XX O53 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 140
XX O54
XX O55 121 PLVQECWVHDCADGK-----LKFCQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRH 172
XX O56 141 PLVQECWVHDCADGKPKSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRH 200
XX O57
XX O58 173 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVVILGRSRLNSNTQGMKFEVENLI 232
XX O59 201 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVVILGRSRLNSNTQGMKFEVENLI 260
XX O60
XX O61 233 LHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCBITGFGK 292
XX O62 261 LHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCBITGFGK 320
XX O63
XX O64 293 ENSTDYLYPEQLKMTVWKLISHRECCQPHYVGVSEVTTKMLCAADPQWKDSCQDSSGGL 352
XX O65 321 ENSTDYLYPEQLKMTVWKLISHRECCQPHYVGVSEVTTKMLCAADPQWKDSCQDSSGGL 380
XX O66
XX O67 353 VCSLQGRMTLTGIVSWGRCALKDKPGYVTVRSHPFLPWRSHKTEENGLAL 403
XX O68 381 VCSLQGRMTLTGIVSWGRCALKDKPGYVTVRSHPFLPWRSHKTEENGLAL 431

RESULT 34
AAP94764
ID AAP94764 standard; protein; 431 AA.
XX
AC AAP94764;
XX
DT 25-MAR-2003 (revised)
DT 27-JUN-1990 (first entry)
XX
DE Non-glycosylated prourokinase.
XX
KW Prourokinase; CGE 195; plasminogen activator; blood clot lysis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Region 1..20
FT /label= signal sequence
XX
PN EP299706-A.
XX
XX 18-JAN-1989.
XX
PF 12-JUL-1988; 88EP-00306334.
XX
PR 13-JUL-1987; 87US-00072426.
PR 29-JUN-1989; 88US-00211279.
XX
PA (COLB) COLLABORATIVE RES INC.
XX
PI Baltimore D, Moir DT, Broeze RJ;
XX
XX WPI; 1989-017204/03.
XX N-PSDB; AAN93079.
XX
XX New non glycosylated, secreted plasminogen activator - pref. with
XX asparagine replaced or deleted, useful for treating blood clots,
XX expressed in non mammalian cells.
XX
XX Disclosure; Page; 26pp; English.
XX
XX myocardial infarction. DNA encoding the protein was sequenced from
XX plasmid pCE195, a subclone of two inserts isolated by screening a cDNA
XX library prep'd. from kidney cell RNA. One of the original inserts, clone
XX CGF31 (tag c) started in the middle of the signal sequence. Mutants of
XX the sequence, pref. in which gcc (Ala) replaces aat (Asn) at nucleotides
XX 1002-1004 (residue 302) are used to transform hosts for the prodn. of non
XX -glycosylated prourokinase. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 431 AA;
XX
XX Query Match 99.2%; Score 2240; DB 1; Length 431;
XX Best Local Similarity 97.8%; Pred. No. 2.1e-174;
XX Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;
XX
XX 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHYRG 60
XX 21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHYRG 80
XX
XX 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
XX 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 140
XX
XX 121 PLVQECWVHDCADGK-----LKFCQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRH 172
XX 141 PLVQECWVHDCADGKPKSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRH 200
XX
XX 173 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVVILGRSRLNSNTQGMKFEVENLI 232
XX 201 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVVILGRSRLNSNTQGMKFEVENLI 260
XX
XX 233 LHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCBITGFGK 292

261 LHKDYSADTLAHNDIALLKIRSEKCAQPSRTIQIICLPMSYNDPQFTGCEITGFGK 320

293 ENSTDYLYPEQLKMTVVKLISHRECQOPHYVGSEVTTKMLCAADPQWKTDSCQDSGGPL 352

321 ENSTDYLYPEQLKMTVVKLISHRECQOPHYVGSEVTTKMLCAADPQWKTDSCQDSGGPL 380

353 VCSLQGRMTLTGIVSMGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 403

381 LCSLQGRMTLTGIVSMGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

ULT 35

37128

AAE37128 standard; protein; 431 AA.

AAE37128;

07-AUG-2003 (first entry)

Human urokinase-type plasminogen activator (uPA) protein.

Osteoarthritis; rheumatoid arthritis; plasmin; plasminogen; human; urokinase-type plasminogen activator; uPA; degenerative joint disease; spondyloarthritis; antisense-therapy; antibody therapy; osteoarthritis; urokinase-type plasminogen activator receptor; psoriatic arthritis; plasminogen-activator inhibitor type 1; PAI-1; uPAR.

Homo sapiens.

WO2003033009-A2.

24-APR-2003.

10-JUL-2002; 2002WO-IB005797.

10-JUL-2001; 2001US-0304461P.

10-JUL-2001; 2001US-030490P.

13-JUL-2001; 2001US-0305182P.

(OMNI-) OMNIO AB.

Ny T, Holmdahl R, Li J;

WPI; 2003-393477/37.

N-PSDB; AAD56133.

Treating or preventing arthritis e.g. osteoarthritis, rheumatoid arthritis in mammals, by administering inhibitors of plasmin, plasminogen, urokinase-type plasminogen-activator or plasminogen-activator inhibitor type 1.

Disclosure; Page 71-73; 85pp; English.

The invention relates to a method for treating or preventing arthritis e.g. osteoarthritis, rheumatoid arthritis in mammals, by administering inhibitors of plasmin, plasminogen, urokinase-type plasminogen activator (uPA), plasminogen-activator inhibitor type 1 (PAI-1) and urokinase-type plasminogen activator receptor (uPAR). The method is useful for treating or preventing arthritis caused by degenerative joint disease, preferably rheumatoid arthritis, psoriatic arthritis, infectious arthritis, juvenile rheumatoid arthritis, osteoarthritis and spondyloarthritis in a mammal, especially a human. It is also useful for identifying agents for treating or preventing arthritis in a mammal and it is also useful in antisense-therapy and antibody therapy. The present sequence is human uPA protein. Note: This sequence is said to be encoded by SEQ ID NO:3 (AAD56133), but this does not appear to be the case

Sequence 431 AA;

Very Match 99.2%; Score 2240; DB 6; Length 431;

est Local Similarity 97.8%; Pred. No. 2.1e-174;

atches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1

Qy	1	SNELHVPNSCDCLNGGTCVSNKYFNSIHWNCNPKKFGQHCBIIDKSKTCYEGNGHFYRG	60
Db	21	SNELHVPNSCDCLNGGTCVSNKYFNSIHWNCNPKKFGQHCBIIDKSKTCYEGNGHFYRG	80
Qy	61	KASTDTMGPRCLPWNSATVLQOITYHAHRSDALQGLGKKNYCRNPNRRRPPWCVVQVGLK	120
Db	81	KASTDTMGPRCLPWNSATVLQOITYHAHRSDALQGLGKKNYCRNPNRRRPPWCVVQVGLK	140
Qy	121	PLVQECVNHDCADGK-----LKPCQCGKTLRPRFKIIGGFEFTTIENQPFAAIYRRH	172
Db	141	PLVQECVNHDCADGKPPSSPEELKFCQCGKTLRPRFKIIGGFEFTTIENQPFAAIYRRH	200
Qy	173	RGGSVTYVCGGSLSPCWISATHCIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLI	232
Db	201	RGGSVTYVCGGSLSPCWISATHCIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLI	260
Qy	233	LHKDYGADTLAHNDIALLKIRSKGRCQAPSTIQTICLPSMYNDPQFTSCITGFGK	292
Db	261	LHKDYGADTLAHNDIALLKIRSKGRCQAPSTIQTICLPSMYNDPQFTSCITGFGK	320
Qy	293	ENSTDVLYPEOLKWTVVVKLISHRECOQPHYVGSEVTTKMLCAADPQWKTDSCQDSDGGPL	352
Db	321	ENSTDVLYPEOLKWTVVVKLISHRECOQPHYVGSEVTTKMLCAADPQWKTDSCQDSDGGPL	380
Qy	353	VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL	403
Db	381	VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL	431
RESULT 36			
ADD46429			
ID	ADD46429 standard; protein; 431 AA.		
XX	AC	ADD46429;	
XX	AC	29-JAN-2004 (first entry)	
DT	Human Protein P00749, SEQ ID NO 12109.		
DE	Human; pain; neuronal tissue; gene therapy;		
XX	KW	spinal segmental nerve injury; chronic constriction injury; CCI;	
KW	KW	spared nerve injury; SNI; Chung.	
XX	XX	Homo sapiens.	
OS			
XX	XX	W02003016475-A2.	
PN	XX	27-FEB-2003.	
PD	XX	14-AUG-2002; 2002WO-US025765.	
PF	XX	14-AUG-2001; 2001US-0312147P.	
XX	PR	01-NOV-2001; 2001US-0346382P.	
PR	PR	26-NOV-2001; 2001US-0333347P.	
XX	XX	(GEHO) GEN HOSPITAL CORP.	
PA	PA	(FARB) BAYER AG.	
XX	XX	Woolf C, D'urso D, Befort K, Costigan M;	
PI	XX	WPI; 2003-268312/26.	
XX	DR	GENBANK; P00749.	
DR	XX	New composition comprising two or more isolated polypeptides, useful for	
PT	PT	preparing a medicament for treating pain in an animal.	
XX	XX	Claim 1; Page; 1017pp; English.	
PS	XX	The invention discloses a composition comprising two or more isolated rat	
XX	CC	or human polynucleotides or a polynucleotide which represents a fragment,	
CC	CC	derivative or allelic variation of the nucleic acid sequence. Also	
CC	CC	claimed are a vector comprising the novel polynucleotide, a host cell	

comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 431 AA;
Query Match 99.2%; Score 2240; DB 7; Length 431;
Best Local Similarity 97.8%; Pred. No. 2.1e-174; Indels 8; Gaps 1;
Matches 402; Conservative 1; Mismatches 0;
1 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60
21 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 80
61 KASDTMTGRPCLPNWSATVLOOTYHAHSDALQGLGKHNYCRPNDRRRPWCYVQVGLK 120
81 KASDTMTGRPCLPNWSATVLOOTYHAHSDALQGLGKHNYCRPNDRRRPWCYVQVGLK 140
121 PLVQECMVHDCADGK-----LKFCQGGKTLRPFKIIIGGEFTTIENQFWFAIYRRH 172
141 PLVQECMVHDCADGKPPSSPPEELKFCQGGKTLRPFKIIIGGEFTTIENQFWFAIYRRH 200
173 RGGSVTYVCGSLSPCWVISATHCFIDYPKEDYIVYLGSRNSNTQGMKEFEVENLI 232
201 RGGSVTYVCGSLSPCWVISATHCFIDYPKEDYIVYLGSRNSNTQGMKEFEVENLI 260
233 LHKDYSADTLAHNDIALLKIRSKEGRCAPQSRITQITCLPSMYNDPQFGTSCBITGFGK 292
261 LHKDYSADTLAHNDIALLKIRSKEGRCAPQSRITQITCLPSMYNDPQFGTSCBITGFGK 320
293 ENSTDYLYPEQLKMTVVKLIHRECOQPHYVGVSEVTTKMLCAADPOWKTSCQDGGGGL 352
321 ENSTDYLYPEQLKMTVVKLIHRECOQPHYVGVSEVTTKMLCAADPOWKTSCQDGGGGL 380
353 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
381 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RE IT 37
AF 4584
II AAR34584 standard; protein; 411 AA.
XZ AAR34584;
XZ
DZ 25-MAR-2003 (revised)
DZ 14-SEP-1993 (first entry)
XZ
XZ Mutant human prourokinase.
XZ PUK; increased half life; improved fibrin affinity.
XZ Homo sapiens.

XX EP541952-A1.
PN 19-MAY-1993.
XX 06-OCT-1992; 92EP-00117000.
XX 07-OCT-1991; 91JP-00289257.
XX (GREC) GREEN CROSS CORP.
XX Tanabe T, Morita M, Hirose M, Amatsuji Y;
XX WPI; 1993-160551/20.
XX N-PSDB; AAQ41450.
XX New human pro-urokinase mutants with thrombolytic activity - have a neutral amino acid in the epidermal growth factor region replaced with a basic amino acid, or an acid residue replaced with a non-acidic residue.
XX Claim 1; Page 17-20; 38pp; English.
XX The sequence is that of a mutant human prourokinase, in which a neutral amino acid in the epidermal growth region has been replaced with a basic amino acid, or an acidic amino acid has been replaced by a non-acidic amino acid. Preferred replacements are 16Gly->Lys, 38Gly->Lys and 45Asp->Asn. The mutant has an increased half-life in blood as compared to the prior art mutant with a deleted EGF region. It has improved affinity for fibrin, and has other features the same as human prourokinase. (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 411 AA;
Query Match 99.2%; Score 2239; DB 2; Length 411;
Best Local Similarity 97.8%; Pred. No. 2.4e-174; Indels 8; Gaps 1;
Matches 402; Conservative 1; Mismatches 0;
1 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60
1 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60
61 KASDTMTGRPCLPNWSATVLOOTYHAHSDALQGLGKHNYCRPNDRRRPWCYVQVGLK 120
61 KASDTMTGRPCLPNWSATVLOOTYHAHSDALQGLGKHNYCRPNDRRRPWCYVQVGLK 120
121 PLVQECMVHDCADGK-----LKFCQGGKTLRPFKIIIGGEFTTIENQFWFAIYRRH 172
121 PLVQECMVHDCADGKPPSSPPEELKFCQGGKTLRPFKIIIGGEFTTIENQFWFAIYRRH 180
173 RGGSVTYVCGSLSPCWVISATHCFIDYPKEDYIVYLGSRNSNTQGMKEFEVENLI 232
181 RGGSVTYVCGSLSPCWVISATHCFIDYPKEDYIVYLGSRNSNTQGMKEFEVENLI 240
233 LHKDYSADTLAHNDIALLKIRSKEGRCAPQSRITQITCLPSMYNDPQFGTSCBITGFGK 292
241 LHKDYSADTLAHNDIALLKIRSKEGRCAPQSRITQITCLPSMYNDPQFGTSCBITGFGK 300
293 ENSTDYLYPEQLKMTVVKLIHRECOQPHYVGVSEVTTKMLCAADPOWKTSCQDGGGGL 352
301 ENSTDYLYPEQLKMTVVKLIHRECOQPHYVGVSEVTTKMLCAADPOWKTSCQDGGGGL 360
353 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
361 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 38
AAU99230
ID AAU99230 standard; protein; 431 AA.
XX AAU99230;
XX AAU99230;
XX 24-SEP-2002 (first entry)
DT

X I Human plasminogen activator, urokinase (PLAU) variant #2.
X F Human; Plasminogen activator; urokinase; PLAU; cancer; enzyme;
X F cytostatic; serine protease; thrombolytic disorder; isogene;
X F pulmonary embolism; chromosome 10q24-qter; haplotype; genotype; SNP;
X F single nucleotide polymorphism; thrombolytic; gene therapy.
X C Homo sapiens.
X C
X C Key Location/Qualifiers
X F Misc-difference 231
X F /note= "Wild-type Lys substituted by Gln"
X F WQ2002040503-A2.
X F 23-MAY-2002.
X F 14-NOV-2001; 2001WQ-US044001.
X F 17-NOV-2000; 2000US-0249703P.
X F (GENA-) GENA-INC. PHARM INC.
X F Anastasio AE, Bentivegna SC, Koshy B;
X F WPI; 2002-519370/55.
X F Genetic variants of Plasminogen activator. Urokinase (PLAU) isogenes,
X F useful for improving efficiency and reliability in drug development for
X F treating thrombolytic disorders and cancer.
X F Claim 27; Page; 92pp; English.
X C The invention relates to a polynucleotide comprising a first nucleotide
X C sequence (NSI) comprising a PLAU (plasminogen activator, urokinase, a
X C serine protease) isogene selected from isogenes 1-9 and 11-20 given in
X C the specification, where each isogene comprises the regions of the PLAU
X C gene or cDNA and is further defined by the corresponding sequence of
X C polymorphisms (defining single nucleotide polymorphisms, SNP). Also
X C included are methods of haplotyping/genotyping (and predicting the
X C haplotype/genotype of the PLAU gene of an individual, identifying an
X C association between a trait and at least one haplotype or haplotype pair
X C of the PLAU gene, an isolated oligonucleotide for detecting a
X C polymorphism in the PLAU gene, a recombinant non-human organism
X C transformed or transfected with the gene or cDNA, fragments of the
X C polynucleotides of at least 10 base pairs encompassing a polymorphic
X C site, an isolated polymorphic variant PLAU protein or fragment, an
X C isolated monoclonal antibody specific for PLAU, a computer system for
X C storing and analysing polymorphism data for the PLAU gene and a genome
X C anchoring for the PLAU gene. PLAU is useful in screening for drugs
X C targeting PLAU that are useful for treating thrombolytic disorders and
X C cancers. The methods are useful for improving the efficiency and
X C reliability of the discovery and development of drugs for treating
X C diseases associated with PLAU activity, in validating PLAU as a drug
X C target and in the design of clinical trials for treating a specific
X C condition of disease associated with PLAU activity. The antibody is
X C useful in diagnostic, prognostic and therapeutic methods. PLAU
X C polynucleotides are useful in studying the expression and function of
X C PLAU, and in expressing PLAU protein for use in screening for candidate
X C drugs to treat diseases related to PLAU activity. The gene for PLAU is
X C located on chromosome 10q24-qter. The present sequence represents a
X C polymorphic variant of the PLAU protein. Note: The present sequence is
X C not shown in the specification but was created by the indexer using the
X C wild-type PLAU protein appearing as AAU99228 and the information on page
X C 28
X C
X C Sequence 431 AA;
X C
X C Query Match 99.2%; Score 2239; DB 5; Length 431;
X C Best Local Similarity 97.8%; Pred. No. 2.5e-174;
X C Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;
X C

QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 80
QY 61 KASDTNMGRCPLPWSNATVLCQTYHAHRSALQIGLGHNYCRNPNRRRPWCYVQVGLK 120
DB 81 KASDTNMGRCPLPWSNATVLCQTYHAHRSALQIGLGHNYCRNPNRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFOGOGKTLRPRFKIIGGFTTITENOPWPAAIYRRH 172
DB 141 PLVQECMVHDCADGKFPSSPEELKFOGOGKTLRPRFKIIGGFTTITENOPWPAAIYRRH 200
QY 173 RGSVTVVCGGSLISPCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
DB 201 RGSVTVVCGGSLISPCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHNDIALIKIRSKGECAPSPRTIOTICLPSMYNDPQFGTSCBITGFGK 292
DB 261 LHKDYSADTLAHNDIALIKIRSKGECAPSPRTIOTICLPSMYNDPQFGTSCBITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVVKLISHRECOQPHYVYGVSEVTTMLCAADPOWKTSCQDGGPL 352
DB 321 ENSTDYLYPEQLKMTVVVKLISHRECOQPHYVYGVSEVTTMLCAADPOWKTSCQDGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTVSVSHFLPWIRSHTKKEENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTVSVSHFLPWIRSHTKKEENGLAL 431

RESULT 39
AAP96146
ID AAP96146 standard; protein; 411 AA.
XX
AC AAP96146;
XX
DT 03-OCT-2002 (revised)
DT 21-JAN-1991 (first entry)
XX
XX Sequence encoded by entire prourokinase (PKU) gene from PKU-producing
DE tumour cell line ATCC CCL138 clone PUC20.
XX
XX Thrombosis; fibrinolytic agent; venous disease; arterial disease therapy.
OS Unidentified.
XX
XX EP312941-A.
XX 26-APR-1989.
XX
XX 15-OCT-1988; 88EP-00117186.
XX
XX 23-OCT-1987; 87DE-03735917.
XX (BADI) BASF AG.
XX Koerwer W, Kurfurst M, Baldinger V, Doerper T, Schwarz M;
PI WPI; 1989-123847/17.
XX N-PSDB; AAN91617.
XX
XX New N-shortened pro-urokinase peptide cpds. with thrombolytic activity -
PT and longer in vivo half life, opt. with replacement of arginine-156.
XX
XX Example; Fig 2, p 705-8; Zipp; German.
XX
XX New N-shortened pro-urokinase peptide cpds. were prepd. from pUC20. The
CC new peptides are useful for treating venous and arterial occlusive
CC diseases. (Updated on 03-OCT-2002 to add missing OS field.)
XX
XX Sequence 411 AA;
XX
XX Query Match 99.2%; Score 2238; DB 1; Length 411;
XX Best Local Similarity 97.8%; Pred. No. 2.8e-174;

	atches	402;	Conservative	0;	Mismatches	1;	Indels	8;	Gaps	1;
Qy	1	SNELHQPNSC	DLNGT	CVSNKYFSNIHWCNCPKFGG	OHCEIDKSKTCYEGNGHFYRG	60				
Df	1	SNELHQPNSC	DLNGT	CVSNKYFSNIHWCNCPKFGG	OHCEIDKSKTCYEGNGHFYRG	60				
Qy	61	KASDTMGRP	CLPWSN	ATVLQOTYHAHRS	DALQGLGKHNYCRNP	PNRRPWCYVQVGLK	120			
Df	61	KASDTMGRP	CLPWSN	ATVLQOTYHAHRS	DALQGLGKHNYCRNP	PNRRPWCYVQVGLK	120			
Qy	121	PLVQECMHD	CADGK	-----LKFCQ	GQKTLRPRFKIIGG	FTTIEQPFPAALYRHH	172			
Df	121	PLVQECMHD	CADGK	-----LKFCQ	GQKTLRPRFKIIGG	FTTIEQPFPAALYRHH	180			
Qy	173	RGGSVTYV	CGSLISPC	WVISA	THCFIDYPKKEDIIV	YVILGRSLNSNTQ	CEMKFEVENLI	232		
Df	181	RGGSVTYV	CGSLISPC	WVISA	THCFIDYPKKEDIIV	YVILGRSLNSNTQ	CEMKFEVENLI	240		
Qy	233	LHKDYSAD	TLAHNDI	ALLKIRSE	KGRCQAPSR	TIOICLP	SMYNDPQGTSC	CEITGFGK	292	
Df	241	LHKDYSAD	TLAHNDI	ALLKIRSE	KGRCQAPSR	TIOICLP	SMYNDPQGTSC	CEITGFGK	300	
Qy	293	ENSTDYLY	PEQLKMT	VVVKLI	SHRECOQPHY	YGVSEVTTKML	CAADPQWK	TDSCQD	SGGPL	352
Df	301	ENSTDYLY	PEQLKMT	VVVKLI	SHRECOQPHY	YGVSEVTTKML	CAADPQWK	TDSCQD	SGGPL	360
Qy	353	VCSLQGR	MTLTC	IVSGW	RGCKALDK	PGVYTRVSHFL	PWIRSH	TKXENGLAL	403	
Df	361	VCSLQGR	MTLTC	IVSGW	RGCKALDK	PGVYTRVSHFL	PWIRSH	TKXENGLAL	411	
RE	40									
AA	2926									
II	AA	R92926	standard;	protein;	411	AA.				
XX	AA	R92926;								
AC	03-AUG-1996	(first entry)								
DT	Pro-urokinase.									
XX	Pro-urokinase.									
DE	Pro-urokinase; plasminogen activator; fusion drug; drug delivery;									
Kx	platelet; cardiovascular disease; thrombolytic.									
Xy	Homo sapiens.									
OS	Key	Location/Qualifiers								
F1	Region	1..132	/label= A-chain							
F1	Domain	1..45	/label= Growth_factor_domain							
F1	Disulfide-bond	11	/note= "disulfide between Cys11 and Cys19"							
F1	Disulfide-bond	13	/note= "disulfide bond between Cys13 and Cys31"							
F1	Disulfide-bond	33	/note= "disulfide bond between Cys33 and Cys42"							
F1	Domain	46..132	/label= Kringle_domain							
F1	Disulfide-bond	50	/note= "disulfide bond between Cys50 and Cys131"							
F1	Disulfide-bond	71	/note= "disulfide bond between Cys71 and Cys113"							
F1	Disulfide-bond	102	/note= "disulfide bond between Cys102 and Cys126"							
F1	Region	133..158	/label= Linker_region							
F1	Disulfide-bond	148	/note= "disulfide bond between Cys148 and Cys273"							
F1	Cleavage-site	156..157	/note= "thrombin cleavage site"							
F1	Cleavage-site	159..159	/note= "plasmin cleavage site"							

FT	Region	159..411								
FT	Disulfide-bond	189	/label= B-chain							
FT	Disulfide-bond	197	/note= "disulfide bond between Cys189 and Cys205"							
FT	Disulfide-bond	293	/note= "disulfide bond between Cys197 and Cys268"							
FT	Disulfide-bond	325	/note= "disulfide bond between Cys293 and Cys362"							
FT	Disulfide-bond	352	/note= "disulfide bond between Cys325 and Cys341"							
FT	Disulfide-bond	352	/note= "disulfide bond between Cys352 and Cys380"							
XX	WO9604004-A1.									
XX	15-FEB-1996.									
XX	03-AUG-1995;	95WO-US009848.								
XX	05-AUG-1994;	94US-00286748.								
XX	(NEW-) NEW ENGLAND DEACONESS HOSPITAL.									
XX	Gurewich V;									
XX	WPI: 1996-129123/13.									
XX	N-PSDB; AAT18237.									
XX	Fusion product of plasminogen activator A chain and drug - targeted to platelets, useful for treatment of cardiovascular disease.									
XX	Claim 3; Page 39-40; 61pp; English.									
XX	A portion of the A-chain (pref. amino acids 1-132) of pro-urokinase (AAR92926) can be linked to a drug, e.g. hirudin, somatostatin or their analogues (see also AAR92925 and AAR92927-33) and used in the prodn. of new fusion drugs. The constructs can be obtcd. by expression of the CC appropriate nucleotide sequences in transformed host cells. When CC administered to a patient, the A-chain binds the fusion drug to the CC platelet outer membrane, i.e. to the site of thrombosis or vascular CC injury. Cleavage sites for thrombin and/or plasmin with the fusion drug CC allow the release of the drug at the target site									
XX	Sequence 411 AA;									
Qy	Query Match	99.2%;	Score 2238;	DB 2;	Length 411;					
Df	Best Local Similarity	97.8%;	Pred. No. 2.8e-174;							
Qy	Matches 402;	Conservative	0;	Mismatches	1;	Indels	8;	Gaps	1;	
Df	1	SNELHQPNSC	DLNGT	CVSNKYFSNIHWCNCPKFGG	OHCEIDKSKTCYEGNGHFYRG	60				
Qy	61	KASDTMGRP	CLPWSN	ATVLQOTYHAHRS	DALQGLGKHNYCRNP	PNRRPWCYVQVGLK	120			
Df	61	KASDTMGRP	CLPWSN	ATVLQOTYHAHRS	DALQGLGKHNYCRNP	PNRRPWCYVQVGLK	120			
Qy	121	PLVQECMHD	CADGK	-----LKFCQ	GQKTLRPRFKIIGG	FTTIEQPFPAALYRHH	172			
Df	121	PLVQECMHD	CADGK	-----LKFCQ	GQKTLRPRFKIIGG	FTTIEQPFPAALYRHH	180			
Qy	173	RGGSVTYV	CGSLISPC	WVISA	THCFIDYPKKEDIIV	YVILGRSLNSNTQ	CEMKFEVENLI	232		
Df	181	RGGSVTYV	CGSLISPC	WVISA	THCFIDYPKKEDIIV	YVILGRSLNSNTQ	CEMKFEVENLI	240		
Qy	233	LHKDYSAD	TLAHNDI	ALLKIRSE	KGRCQAPSR	TIOICLP	SMYNDPQGTSC	CEITGFGK	292	
Df	241	LHKDYSAD	TLAHNDI	ALLKIRSE	KGRCQAPSR	TIOICLP	SMYNDPQGTSC	CEITGFGK	300	
Qy	293	ENSTDYLY	PEQLKMT	VVVKLI	SHRECOQPHY	YGVSEVTTKML	CAADPQWK	TDSCQD	SGGPL	352
Df	301	ENSTDYLY	PEQLKMT	VVVKLI	SHRECOQPHY	YGVSEVTTKML	CAADPQWK	TDSCQD	SGGPL	360

353 VCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHTKBENG LAL 403
361 VCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHTKBENG LAL 411

ULT 41
74797
AAB74797 standard; protein; 411 AA.
AAB74797;
12-JUN-2001 (first entry)
Prourokinase protein sequence.
Prourokinase, Pro-309; mutagenic; urokinase; zymogen; mutant;
lowered fibrinogen dissolving activity; fibrin; E segment; D segment;
lowered non-specific fibrin dissolving zymogen activation.
Unidentified.

Key Location/Qualifiers
Disulfide-bond 11..19
Disulfide-bond 13..31
Disulfide-bond 33..42
Disulfide-bond 50..131
Disulfide-bond 71..113
Disulfide-bond 102..126
Disulfide-bond 148..279
Disulfide-bond 189..205
Disulfide-bond 197..268
Disulfide-bond 293..362
Disulfide-bond 325..341
Disulfide-bond 352..380

CN1277262-A.
20-DEC-2000.
10-JUL-2000; 2000CN-00109829.
10-JUL-2000; 2000CN-00109829.
(LIUJ/) LIU J.
Sun Z, Liu J;
WPI; 2001-266614/28.
Urokinase zymogen mutant.
Example; Fig.1; 1lpp; Chinese.
The present invention describes a prourokinase mutant comprising the amino acid sequence point mutation at proline 309. The mutation makes the mutant have an intrinsic activity 2.5-20 times lower than that of natural prourokinase, including lowered fibrinogen dissolving activity and lowered non-specific fibrin dissolving zymogen activation. Compared with the natural prourokinase, the fibrin dissolving zymogen activation of the mutant may be promoted by not only the E segment of degraded fibrin but also the D segment. The present sequence represents a wild type prourokinase.protein sequence which is used in an example from the present invention. N.B. The sequence in the specification is of poor quality so the sequence given here is of the indexers best interpretation

Sequence 411. AA;
Very Match 99.2%; Score 2238; DB 4; Length 411;
set Local Similarity 97.8%; Pred. No. 2.8e-174;
atches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;
1 SNEHVPNSDCLNGTCTVSNKYFNIHWCNCFKFGGHCIDKSKTCYEGNGHYRG 60

Db 1 SNEHVPNSDCLNGTCTVSNKYFNIHWCNCFKFGGHCIDKSKTCYEGNGHYRG 60
Qy 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRPNRRPWCYVQVLK 120
Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRPNRRPWCYVQVLK 120
Qy 121 PLVCEWVHDCADGK-----LKFCGOKTLRPRFKIIGGFTTIIENOPWFAAIYRRH 172
Db 121 PLVCEWVHDCADGKFPPEELKFCGOKTLRPRFKIIGGFTTIIENOPWFAAIYRRH 180
Qy 173 RGSVTVVCGSLSPCWISATHCFIDYPKEDYIVYLGSRSLNSNTQEMKFEVENLI 232
Db 181 RGSVTVVCGSLSPCWISATHCFIDYPKEDYIVYLGSRSLNSNTQEMKFEVENLI 240
Qy 233 LHKDYSADTLAHNDIALKIRSKGCAQPSRTIQTICLPSMYNDPQGTSCITGFGK 292
Db 241 LHKDYSADTLAHNDIALKIRSKGCAQPSRTIQTICLPSMYNDPQGTSCITGFGK 300
Qy 293 ENSTDYLYPEOLKMTVVVKLISHRECQPPHYVGSVTTMLCAADPQWKTDSCQDSGGPL 352
Db 301 ENSTDYLYPEOLKMTVVVKLISHRECQPPHYVGSVTTMLCAADPQWKTDSCQDSGGPL 360
Qy 353 VCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHTKBENG LAL 403
Db 361 VCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHTKBENG LAL 411

RESULT 42
AAU99229 standard; protein; 431 AA.
AC AAU99229;
DT 24-SEP-2002 (first entry)
XX Human plasminogen activator, urokinase (PLAU) variant #1.
XX Human; Plasminogen activator; urokinase; PLAU; cancer; enzyme;
KW cytosolic; serine protease; thrombolytic disorder; isogene;
KW pulmonary embolism; chromosome 10q24-qter; haplotype; genotype; SNP;
KW single nucleotide polymorphism; thrombolytic; gene therapy.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Misc-difference 350
FT /note= "Wild-type Tyr substituted by His"
XX
XX WO200240503-A2.
XX 23-MAY-2002.
XX 14-NOV-2001; 2001WO-US044001.
XX 17-NOV-2000; 2000US-0249703P.
XX (GENA-) GENAISSANCE PHARM INC.
XX Anastasio AE, Bentivegna SC, Koshy B;
XX WPI; 2002-S19370/55.
XX Genetic variants of plasminogen activator, Urokinase (PLAU) isogenes,
PT useful for improving efficiency and reliability in drug development for
PT treating thrombolytic disorders and cancer.
XX Claim 27; Page; 92pp; English.
XX The invention relates to a polynucleotide comprising a first nucleotide
CC sequence (NS1) comprising a PLAU (plasminogen activator, urokinase, a
CC serine protease) isogene selected from isogenes 1-9 and 11-20 given in
CC the specification, where each isogene comprises the regions of the PLAU
CC gene or cDNA and is further defined by the corresponding sequence of

polymorphisms (defining single nucleotide polymorphisms, SNP). Also included are methods of haplotyping/genotyping (and predicting the haplotype/genotype of the PLAU gene of an individual, identifying an association between a trait and at least one haplotype or haplotype pair of the PLAU gene, an isolated oligonucleotide for detecting a polymorphism in the PLAU gene, a recombinant non-human organism transformed or transfected with the gene or cDNA, fragments of the polynucleotides of at least 10 base pairs encompassing a polymorphic site, an isolated polymorphic variant PLAU protein or fragment, an isolated monoclonal antibody specific for PLAU, a computer system for storing and analysing polymorphism data for the PLAU gene and a genome targeting PLAU that are useful for screening for drugs and an antibody for PLAU that are useful for treating thrombotic disorders and cancers. The methods are useful for improving the efficiency and reliability of the discovery and development of drugs for treating diseases associated with PLAU activity, in validating PLAU as a drug target and in the design of clinical trials for treating a specific condition of disease associated with PLAU activity. The antibody is useful in diagnostic, prognostic and therapeutic methods. PLAU polynucleotides are useful in studying the expression and function of PLAU, and in expressing PLAU protein for use in screening for candidate drugs to treat diseases related to PLAU activity. The gene for PLAU is located on chromosome 10q24-qter. The present sequence represents a polymorphic variant of the PLAU protein. Note: The present sequence is not shown in the specification but was created by the indexer using the wild-type PLAU protein appearing as AAU99228 and the information on page 28

Sequence 431 AA;

Query Match 99.2%; Score 2238; DB 5; Length 431;
 Best Local Similarity 97.8%; Pred. No. 3e-174;
 Mismatches 1; Indels 8; Gaps 1;
 Conservative 1; Mismatches 0; Indels 8; Gaps 1;
 1 SNELHQPSPNCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
 21 SNELHQPSPNCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 80
 61 KASDTMGRPCLPWNASVTLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
 81 KASDTMGRPCLPWNASVTLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPWCYVQVGLK 140
 121 PLVQECWVHDCADGK-----LKFCQCGQKTLPRFKIIGGFTTIIENQWFAAIYRRH 172
 141 PLVQECWVHDCADGKPKSPPEELKFCQCGQKTLPRFKIIGGFTTIIENQWFAAIYRRH 200
 173 RGSVTVYVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 232
 201 RGSVTVYVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 260
 233 LHKDYSADTLAHHNDIALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFTSCBITGFGK 292
 261 LHKDYSADTLAHHNDIALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFTSCBITGFGK 320
 293 ENSTDYLYPEQLKMTVVKLISHRECQPHYVGSVTTKWLCAADPQWKTDCSQGDSGGL 352
 321 ENSTDYLYPEQLKMTVVKLISHRECQPHYVGSVTTKWLCAADPQWKTDCSQGDSGGL 380
 353 VCSLQGRMTLTGIVSGWRCALCKDPGYVTVRSVHFLPWRSHTKKEENGAL 403
 381 VCSLQGRMTLTGIVSGWRCALCKDPGYVTVRSVHFLPWRSHTKKEENGAL 431

UT 43
 2999

AAR62999 standard; protein; 411 AA.

AAR62999;

25-MAR-2003 (revised)

21-SEP-1995 (first entry)

Pro-urokinase mutant His613.

XX Pro-urokinase; thrombolysis; fibrin clot lysis; mutant His13;
 KW reduced fibrinogenolysis; non-specific plasminogen activation;
 XX systemic bleeding.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FT Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Domain 297..313 /note= "flexible loop"
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 PN WO9501427-A1.
 XX 12-JAN-1995.
 XX 28-JUN-1994; 94WO-US007278.
 XX 02-JUL-1993; 93US-00087163.
 PA (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Liu J, Gurewich V;
 XX WPI; 1995-060991/08.
 XX Pro-urokinase mutants - have thrombolytic activity but reduced
 PT fibrinogenolysis activity and non-specific plasminogen activation.
 XX Claim 11; Fig 1; 46pp; English.
 XX AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the treatment
 CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
 CC specific plasminogen activation. The mutants can therefore be used for
 CC the lysis of fibrin clots without inducing systemic bleeding, as can be
 CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
 CC PN field.)
 XX SQ Sequence 411 AA;
 Query Match 99.1%; Score 2237; DB 2; Length 411;
 Best Local Similarity 97.8%; Pred. No. 3.4e-174;
 Matches 402; Conservative 0; Mismatches 1; Indels 8; Gaps 1;
 QY 1 SNELHQPSPNCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
 DB 1 SNELHQPSPNCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
 QY 61 KASDTMGRPCLPWNASVTLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
 DB 61 KASDTMGRPCLPWNASVTLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
 QY 121 PLVQECWVHDCADGK-----LKFCQCGQKTLPRFKIIGGFTTIIENQWFAAIYRRH 172
 DB 121 PLVQECWVHDCADGKPKSPPEELKFCQCGQKTLPRFKIIGGFTTIIENQWFAAIYRRH 180
 QY 173 RGSVTVYVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 232
 DB 181 RGSVTVYVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 240

233 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIOTICLPMSYNDPQFGTSCETGFGK 292
241 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIOTICLPMSYNDPQFGTSCETGFGK 300
293 ENSTDYLYPEQLKMTVVVKLI SHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGFL 352
301 ENSTDYLYPEQLKMTVVVKLI SHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGFL 360
353 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 403
361 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 411

JUL 44
62992
AAR62992 standard; protein; 411 AA.
AAR62992;
25-MAR-2003 (revised)
21-SEP-1995 (first entry)
Pro-urokinase mutant Ala300.
Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala300;
reduced fibrinogenolysis; non-specific plasminogen activation;
systemic bleeding.
Homo sapiens.
Key Location/Qualifiers
Disulfide-bond 11..19
Disulfide-bond 13..31
Disulfide-bond 33..42
Disulfide-bond 50..131
Disulfide-bond 71..113
Disulfide-bond 102..126
Disulfide-bond 148..279
Disulfide-bond 189..205
Disulfide-bond 197..268
Disulfide-bond 293..362
Domain 297..313
Disulfide-bond 325..341
Disulfide-bond 352..380
WO9501427-A1.
12-JAN-1995.
28-JUN-1994; 94WO-US007278.
02-JUL-1993; 93US-00087163.
(NEW-) NEW ENGLAND DEACONESS HOSPITAL.
Liu J, Gurewich V;
WPI; 1995-060991/08.
Pro-urokinase mutants - have thrombolytic activity but reduced
fibrinogenolysis activity and non-specific plasminogen activation.
Claim 5; Fig 1; 46pp; English.
AAR62991 is the wild type pro-urokinase, from which the new mutants
described in AAR62992-R63008 were derived. These mutants retain the
thrombolytic activity of the wild type protein, useful for the treatment
of thromboembolism, but have a reduced fibrinogenolysis activity and non-
specific plasminogen activation. The mutants can therefore be used for
the lysis of fibrin clots without inducing systemic bleeding, as can be
the case with the wild type protein. (Updated on 25-MAR-2003 to correct
PN field.)

XX Sequence 411 AA;
SQ
Query Match 99.1%; Score 2237; DB 2; Length 411;
Best Local Similarity 97.8%; Pred. No. 3.4e-174;
Matches 402; Conservative 0; Mismatches 1; Indels 8; Gaps 1;
QY 1 SNEHGVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNEHGVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQGLK 120
DB 61 KASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQGLK 120
QY 121 PLVQECMVHDCADGK-----LKFCQGGKTLPRPKLIIGGEFTTIENQPFAAIYRRH 172
DB 121 PLVQECMVHDCADGKPPSEELKFCQGGKTLPRPKLIIGGEFTTIENQPFAAIYRRH 180
QY 173 RGSVTVYVCGGSLISPCWVISATHCIFIDYPKKEDYIVVLGRSRLNSNTQGEKKEVENLI 232
DB 181 RGSVTVYVCGGSLISPCWVISATHCIFIDYPKKEDYIVVLGRSRLNSNTQGEKKEVENLI 240
QY 233 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIOTICLPMSYNDPQFGTSCETGFGK 292
DB 241 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIOTICLPMSYNDPQFGTSCETGFGK 300
QY 293 ENSTDYLYPEQLKMTVVVKLI SHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGFL 352
DB 301 ENSTDYLYPEQLKMTVVVKLI SHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGFL 360
QY 353 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 361 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 45
AAR62998
ID AAR62998 standard; protein; 411 AA.
AC AAR62998;
DT 25-MAR-2003 (revised)
DT 21-SEP-1995 (first entry)
XX
DE Pro-urokinase mutant Ala313.
XX
KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala313;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
systemic bleeding.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Domain 297..313
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX PD 12-JAN-1995.
XX

PF 28-JUN-1994; 94WO-US007278.
XX
XX
XX 02-JUL-1993; 93US-00087163.
XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
XX Liu J, Gurewich V;
XX WPI; 1995-060991/08.
XX
XX Pro-urokinase mutants - have thrombolytic activity but reduced
XX fibrinogenolysis activity and non-specific plasminogen activation.
XX
XX Claim 11; Fig 1; 46pp; English.
XX
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
XX described in AAR62992-R63008 were derived. These mutants retain the
XX thrombolytic activity of the wild type protein, useful for the treatment
XX of thromboembolism, but have a reduced fibrinogenolysis activity and non-
XX specific plasminogen activation. The mutants can therefore be used for
XX the lysis of fibrin clots without inducing systemic bleeding, as can be
XX the case with the wild type protein. (Updated on 25-MAR-2003 to correct
XX PN field.)
XX
XX Sequence 411 AA;
XX
XX Query Match 99.1%; Score 2237; DB 2; Length 411;
XX St Local Similarity 97.8%; Pred. No. 3.4e-174;
XX Matches 402; Conservative 0; Mismatches 1; Indels 8; Gaps 1;
XX
XX 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
XX 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
XX
XX 61 KASDTMTGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
XX 61 KASDTMTGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
XX
XX 121 PLVQECMWHDCADGK-----LKFCGQKTLRPRFKIIGGEPTTIENQFWFAAIYRRH 172
XX 121 PLVQECMWHDCADGKPPPEELKFCGQKTLRPRFKIIGGEPTTIENQFWFAAIYRRH 180
XX
XX 173 RGGSVTYVCGSLSPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 232
XX 181 RGGSVTYVCGSLSPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 240
XX
XX 233 LHKDYSADTLAHHNDIALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
XX 241 LHKDYSADTLAHHNDIALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 300
XX
XX 293 ENSTDYLYPEQLKMTVVKLI SHRCQQPHYVGSVTTKMLCAADPQWKTDSCQDSSGPL 352
XX 301 ENSTDYLYPEQLAMTVVKLI SHRCQQPHYVGSVTTKMLCAADPQWKTDSCQDSSGPL 360
XX
XX 353 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
XX 361 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
XX
XX LT 46
XX AAR62993 standard; protein; 411 AA.
XX
XX AAR62993;
XX
XX 25-MAR-2003 (revised)
XX 21-SEP-1995 (first entry)
XX
XX Pro-urokinase mutant His300.
XX
XX Pro-urokinase; thrombolysis; fibrin clot lysis; mutant His300;
XX reduced fibrinogenolysis; non-specific plasminogen activation;
XX systemic bleeding.

XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Disulfide-bond 11..19
XX Disulfide-bond 13..31
XX Disulfide-bond 33..42
XX Disulfide-bond 50..131
XX Disulfide-bond 71..113
XX Disulfide-bond 102..126
XX Disulfide-bond 148..279
XX Disulfide-bond 189..205
XX Disulfide-bond 197..268
XX Disulfide-bond 293..362
XX Domain /note="flexible loop"
XX Disulfide-bond 325..341
XX Disulfide-bond 352..380
XX WO9501427-A1.
XX
XX 12-JAN-1995.
XX
XX 28-JUN-1994; 94WO-US007278.
XX
XX 02-JUL-1993; 93US-00087163.
XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
XX
XX Liu J, Gurewich V;
XX WPI; 1995-060991/08.
XX
XX Pro-urokinase mutants - have thrombolytic activity but reduced
XX fibrinogenolysis activity and non-specific plasminogen activation.
XX
XX Claim 5; Fig 1; 46pp; English.
XX
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
XX described in AAR62992-R63008 were derived. These mutants retain the
XX thrombolytic activity of the wild type protein, useful for the treatment
XX of thromboembolism, but have a reduced fibrinogenolysis activity and non-
XX specific plasminogen activation. The mutants can therefore be used for
XX the lysis of fibrin clots without inducing systemic bleeding, as can be
XX the case with the wild type protein. (Updated on 25-MAR-2003 to correct
XX PN field.)
XX
XX Sequence 411 AA;
XX
XX Query Match 99.1%; Score 2237; DB 2; Length 411;
XX Best Local Similarity 97.8%; Pred. No. 3.4e-174;
XX Matches 402; Conservative 0; Mismatches 1; Indels 8; Gaps 1;
XX
XX 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
XX 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
XX
XX 61 KASDTMTGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
XX 61 KASDTMTGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
XX
XX 121 PLVQECMWHDCADGK-----LKFCGQKTLRPRFKIIGGEPTTIENQFWFAAIYRRH 172
XX 121 PLVQECMWHDCADGKPPPEELKFCGQKTLRPRFKIIGGEPTTIENQFWFAAIYRRH 180
XX
XX 173 RGGSVTYVCGSLSPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 232
XX 181 RGGSVTYVCGSLSPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 240
XX
XX 233 LHKDYSADTLAHHNDIALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
XX 241 LHKDYSADTLAHHNDIALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGH 300
XX
XX SQ

C 293 ENSTDYLYPEQLKMTVVVKLIHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDSGGPL 352
D 301 ENSTDYLYPEQLKMTVVVKLIHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDSGGPL 360
C 353 VCSLQGRWTLTGIVSWGRCALKDXPGVTVRVSHFLPWIRSHTKENGLAL 403
E 361 VCSLQGRWTLTGIVSWGRCALKDXPGVTVRVSHFLPWIRSHTKENGLAL 411
ULT 47
A 53000
I AAR63000 standard; protein; 411 AA.
X AAR63000;
X
X
X 25-MAR-2003 (revised)
D 21-SEP-1995 (first entry)
X
X
X Pro-urokinase mutant Ser175 His187.
X
X Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;
K non-specific plasminogen activation; systemic bleeding;
K mutant Ser175 His187.
X
X Homo sapiens.
X
X Key Location/Qualifiers
E Disulfide-bond 11..19
E Disulfide-bond 13..31
E Disulfide-bond 33..42
F Disulfide-bond 50..131
E Disulfide-bond 71..113
E Disulfide-bond 102..126
F Disulfide-bond 148..279
F Disulfide-bond 189..205
E Disulfide-bond 197..268
E Disulfide-bond 293..362
F Domain 297..313
F /note= "flexible loop"
F Disulfide-bond 325..341
F Disulfide-bond 352..380
X
X WO9501427-A1
X
X 12-JAN-1995.
X
X 28-JUN-1994; 94WO-US007278.
X
X 02-JUL-1993; 93US-00087163.
X (NEWB-) NEW ENGLAND DEACONESS HOSPITAL.
X
X Liu J, Gurewich V;
X WPI; 1995-060991/08.
X
X Pro-urokinase mutants - have thrombolytic activity but reduced
F fibrinogenolysis activity and non-specific plasminogen activation.
X
X Claim 15; Fig 1; 46pp; English.
X
X AAR62991 is the wild type pro-urokinase, from which the new mutants
C described in AAR62992-R63008 were derived. These mutants retain the
C thrombolytic activity of the wild type protein, useful for the treatment
C of thromboembolism, but have a reduced fibrinogenolysis activity and non-
C specific plasminogen activation. The mutants can therefore be used for
C the lysis of fibrin clots without inducing systemic bleeding, as can be
C the case with the wild type protein. (Updated on 25-MAR-2003 to correct
C PN field.)
X
X Sequence 411 AA;
X
X very Match 99.0%; Score 2235; DB 2; Length 411;

Best Local Similarity 97.6%; Pred. No. 5e-174;
Matches 401; Conservative 2; Mismatches 0; Indels 8; Gaps 1;
Oy 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKXFGGQHCIDKSKTCYEGNGHYRG 60
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKXFGGQHCIDKSKTCYEGNGHYRG 60
Oy 61 KASDTMTGRPCLPWN SATVLQOYTHAHRSDALQGLGKHNYCRNPNRRRPWCYVQVGLK 120
Db 61 KASDTMTGRPCLPWN SATVLQOYTHAHRSDALQGLGKHNYCRNPNRRRPWCYVQVGLK 120
Oy 121 PLVOECMVHDCADGK-----LKFOCGOKTLRPRFKIIGGFETTIENTOPWFAAIYRRH 172
Db 121 PLVOECMVHDCADGKFPPELKFQCGOKTLRPRFKIIGGFETTIENTOPWFAAIYRRH 180
Oy 173 RGSVTVVCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQEMKFEVENLI 232
Db 181 RGSVTVVCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQEMKFEVENLI 240
Oy 233 LHKDYSADTLAHHNDIALALKIRSKGRCAPSRITQICLPWMYNDPQGTSCIEITGFGK 292
Db 241 LHKDYSADTLAHHNDIALALKIRSKGRCAPSRITQICLPWMYNDPQGTSCIEITGFGK 300
Oy 293 ENSTDYLYPEQLKMTVVVKLIHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDSGGPL 352
Db 301 ENSTDYLYPEQLKMTVVVKLIHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDSGGPL 360
Oy 353 VCSLQGRWTLTGIVSWGRCALKDXPGVTVRVSHFLPWIRSHTKENGLAL 403
Db 361 VCSLQGRWTLTGIVSWGRCALKDXPGVTVRVSHFLPWIRSHTKENGLAL 411
RESULT 48
AAY39343
ID AAY39343 standard; protein; 411 AA.
XX
AC AAY39343;
XX
DT 01-DEC-1999 (first entry)
DE Human pro-urokinase.
XX
KW Serine protease; plasminogen; plasmin; activation; matrix; cancer;
KW tumour; metastasis; X-ray crystallography; inhibitor.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Protein 1..158
FT /label= Mature_urokinase_A_chain
FT Protein 1..135
FT /label= Low_molecular_weight_urokinase
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 135..136
FT Cleavage-site /note= "Additional cleavage at this site generates low
FT molecular weight (LMW) urokinase"
FT Disulfide-bond 148..279
FT /note= "Links mature urokinase A- and B-chains"
FT Cleavage-site 158..159
FT /note= "Cleavage at this site generates mature urokinase
FT A- and B-chains"
FT Protein 159..411
FT /label= Mature_urokinase_B_chain
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Modified-site 302
FT /note= "N-glycosylated"
FT

FT Disulfide-bond 325. .341
FT Disulfide-bond 352. .380
XX Cleavage-site 405. .406
PN WO9945379-A2.
XX
PD 10-SEP-1999.
XX
XX 05-MAR-1999; 99WO-US004967.
XX
XX 06-MAR-1998; 98US-00036184.
XX
XX (ABBO) ABBOTT LAB.
XX
XX Nienaber VL, Greer J, Abad-Zapatero C, Norbeck DW;
XX
XX WPI; 1999-571607/48.
XX
XX

PE Identifying ligands for target biomolecules using X-ray crystallography.
XX
XX Example 1; Fig 5; 57pp; English.
XX

XX This sequence represents human pro-urokinase. The mature urokinase
CC consists of an A- and B-chain, linked by a single disulphide bond, and is
CC generated by proteolytic cleavage of the peptide bond between Lys 158 and
CC Ile 159. Additional cleavage of the peptide bond between Lys 135 and Lys
CC Ile 159 generates a low molecular weight urokinase. The urokinase A-chain
CC 136 generates a low molecular weight urokinase. The urokinase A-chain
CC contains an EGF-like domain and a kringle domain, while the B-chain
CC contains the catalytic domain. Urokinase is a serine protease and is
CC strongly associated with tumour cells. Urokinase activates plasminogen
CC which, in turn, activates the matrix metalloproteinases. Plasmin and the
CC metalloproteinases degrade the extracellular matrix and promote tumour
CC growth and metastasis. Inhibitors that specifically target urokinase may
CC serve as effective anticancer agents. A novel method for identifying such
CC ligands used X-ray crystallography to determine if a complex is formed
CC between a ligand and a target biomolecule. However, crystals of a native
CC urokinase/inhibitor complex had poor diffraction quality. Human urokinase
CC was therefore engineered so that it would produce crystals with the
CC desired qualities. This engineered urokinase was designated mu-UK
CC (RAY39344).
XX
XX Sequence 411 AA;

ery Match 99.0%; Score 2235; DB 2; Length 411;
st Local Similarity 97.8%; Pred. No. 5e-174;
tches 402; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

Qy 1 SNEHQPNSCCLNGTGVSNKYFNIHWCNCPKFGQHCIDKSKTCYEGNGHYRG 60
Df 1 SNEHQPNSCCLNGTGVSNKYFNIHWCNCPKFGQHCIDKSKTCYEGNGHYRG 60
Qy 61 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQGLK 120
Df 61 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQGLK 120
Qy 121 PLVQECMVHDCADGK-----LKFCGGQKTLAPREKLIIGGFTTIENQFWFAAIYRRH 172
Df 121 PLVQECMVHDCADGKSPPEEKLPFCQCKTLAPREKLIIGGFTTIENQFWFAAIYRRH 180
Qy 173 RGGSVYVYCGGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
Df 181 RGGSVYVYCGGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLI 240
Qy 233 LHKDYSADTLAHHNDIALKTRSKRGCAQPSRTIQCILPSMYNDPQFTSCBITGFGK 292
Df 241 LHKDYSADTLAHHNDIALKTRSKRGCAQPSRTIQCILPSMYNDPQFTSCBITGFGK 300
Qy 293 ENSDTLYPEQLKMTVVKLISHRECQPPHYGSEVTTNMLCAADPQWKTDSCQDSDGGPL 352
Df 301 ENSDTLYPEQLKMTVVKLISHRECQPPHYGSEVTTNMLCAADPQWKTDSCQDSDGGPL 360
Qy 353 VCSLQGRMTLGTIVSWGRGCAKDKPGVYTVSHFLPWRSHTKKEENGLAL 403

Db 361 VCSLQGRMTLGTIVSWGRGCAKDKPGVYTVSHFLPWRSHTKKEENGLAL 411
RESULT 49
RAY42284
ID AAY42284 standard; protein; 411 AA.
XX
XX AAY42284;
AC
DT 01-DEC-1999 (first entry)
XX
XX Human pro-urokinase.
DE
XX Serine protease; plasminogen; plasmin; activation; matrix; cancer;
KW tumour; metastasis; X-ray crystallography; inhibitor.
KW
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1. .158
FT Protein /label= Mature_urokinase_A_chain
FT Protein 1. .135
FT Protein /label= Low_molecular_weight_urokinase
FT Disulfide-bond 11. .19
FT Disulfide-bond 13. .31
FT Disulfide-bond 33. .42
FT Disulfide-bond 50. .131
FT Disulfide-bond 71. .113
FT Disulfide-bond 102. .126
FT Cleavage-site 135. .136
FT /note= "Additional cleavage at this site generates low
FT molecular weight (LMW) urokinase"
FT Disulfide-bond 148. .279
FT Cleavage-site 158. .159
FT /note= "Links mature urokinase A- and B-chains"
FT /note= "Cleavage at this site generates mature urokinase
FT A- and B-chains"
FT Protein 159. .411
FT /label= Mature_urokinase_B_chain
FT Disulfide-bond 189. .205
FT Disulfide-bond 197. .268
FT Disulfide-bond 293. .362
FT Modified-site 302
FT /note= "N-glycosylated"
FT Disulfide-bond 325. .341
FT Disulfide-bond 352. .380
FT Cleavage-site 405. .406
XX WO9945389-A2.
PN
XX 10-SEP-1999.
PD
XX 01-MAR-1999; 99WO-US004518.
PF
XX 06-MAR-1998; 98US-00036184.
PR
XX (ABBO) ABBOTT LAB.
XX
XX Nienaber VL, Greer J, Abad-Zapatero C, Norbeck DW;
PI
XX WPI; 1999-551079/46.
DR
XX Identifying ligands for target biomolecules using X-ray crystallography,
PT used for designing ligands with improved biological activity for target
PT receptor.
XX
XX Example 1; Fig 5; 57pp; English.
PS
XX This sequence represents human pro-urokinase. The mature urokinase
CC consists of an A- and B-chain, linked by a single disulphide bond, and is
CC generated by proteolytic cleavage of the peptide bond between Lys 158 and
CC Ile 159. Additional cleavage of the peptide bond between Lys 135 and Lys
CC Ile 159 generates a low molecular weight urokinase. The urokinase A-chain
CC 136 generates a low molecular weight urokinase. The urokinase A-chain

contains an EGF-like domain and a kringle domain, while the B-chain contains the catalytic domain. Urokinase is a serine protease and is strongly associated with tumour cells. Urokinase activates plasminogen which, in turn, activates the matrix metalloproteinases. Plasmin and the metalloproteinases degrade the extracellular matrix and promote tumour growth and metastasis. Inhibitors that specifically target urokinase may serve as effective anticancer agents. A novel method for identifying such ligands used X-ray crystallography to determine if a complex is formed between a ligand and a target biomolecule. However, crystals of a native urokinase/inhibitor complex had poor diffraction quality. Human urokinase was therefore engineered so that it would produce crystals with the desired qualities. This engineered urokinase was designated mu-UK (AAV42285)

Sequence 411 AA;

Query Match 99.0%; Score 2235; DB 2; Length 411;
 Best Local Similarity 97.8%; Pred. No. 5e-174;
 Matches 402; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

1 SNELHQPNSCCLNGGTCVSNKYFSNTHWNCNPKKFGQCEIDKSKTCYEGNGHFYRG 60
 1 SNELHQPNSCCLNGGTCVSNKYFSNTHWNCNPKKFGQCEIDKSKTCYEGNGHFYRG 60
 61 KASTDTMGPRCLPWNASATVLQOTYHAHRSALQGLGKHNYCRNPNRRRPPWCYVQGLK 120
 61 KASTDTMGPRCLPWNASATVLQOTYHAHRSALQGLGKHNYCRNPNRRRPPWCYVQGLK 120
 121 PLVQECVHDCADGK-----LKFQCGQKTLRPRFKLIGGEFTTIENQWFAAIVERH 172
 121 PLVQECVHDCADGKPPPELPPQCGQKTLRPRFKLIGGEFTTIENQWFAAIAIRH 180
 173 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGEKMFVENLI 232
 181 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGEKMFVENLI 240
 233 LHKDYSADTLAHNDIALKIRSKGRCAQPERTTQICLPSMYNDPQGSCEITGTGK 292
 241 LHKDYSADTLAHNDIALKIRSKGRCAQPERTTQICLPSMYNDPQGSCEITGTGK 300
 293 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKDSCQGSGGPL 352
 301 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKDSCQGSGGPL 360
 353 VCSLQGRMTLTGVSWGRGCAKDKPGVTVTRVSHFLPWTIRSHTKENGIAL 403
 361 VCSLQGRMTLTGVSWGRGCAKDKPGVTVTRVSHFLPWTIRSHTKENGIAL 411

R ULT 50

A 20489
 AAB20489 standard; protein; 411 AA.

X A

X A

X A

X 21-JUN-2001 (first entry)

X Human pro-urokinase plasminogen activator.

X Urokinase plasminogen activator; uPA; human; tumour; cell migration;
 K cell invasion; cell proliferation; angiogenesis; apoptosis; antitumour;
 K diagnosis; therapy.

X Homo sapiens.

F Key Location/Qualifiers
 F Domain 1. .135
 F /label= ATF
 F /note= "the ATF domain alternatively comprises residues 1-43"

F Domain 4 .43

F Disulfide-bond 11. .19
 F /label= Growth_factor_domain

F Disulfide-bond 11. .19

FT Disulfide-bond 13. .31
 FT Disulfide-bond 33. .42
 FT Disulfide-bond 50. .131
 FT Disulfide-bond 71. .113
 FT Disulfide-bond 102. .126
 FT Disulfide-bond 148. .279
 FT Cleavage-site 158. .159
 FT /note= "cleavage at this site results in the formation of
 FT the two-chain active uPA (tcuPA)"
 FT Disulfide-bond 189. .205
 FT Disulfide-bond 197. .268
 FT Disulfide-bond 293. .362
 FT Disulfide-bond 325. .341
 FT Disulfide-bond 352. .380
 XX
 PN W0200125410-A2.

XX 12-APR-2001.

XX 27-SEP-2000; 2000WO-US026502.

XX 01-OCT-1999; 99US-0157012P.

XX (ANGS-) ANGSTROM PHARM INC.

XX Mazar AP, Jones TR;

XX WPI; 2001-290611/30.

XX Novel urokinase plasminogen activator cell surface receptor-targeting
 PT protein or peptide, useful for inhibiting angiogenesis or cell migration,
 PT invasion or proliferation, is diagnostically or therapeutically labeled.

XX Disclosure; Fig 1; 35pp; English.

XX The present sequence is that of human pro-urokinase plasminogen activator
 CC (pro-uPA). The invention provides a uPA receptor (uPAR) targeting protein
 CC or peptide that is labelled and used in methods of diagnosis and therapy.
 CC The labelled protein or peptide preferably has the following properties:
 CC It comprises at least 38 amino acid residues, including residues 13-30 of
 CC the uPAR binding site of uPA; competes with labelled BFP-uPA for binding
 CC to a cell or molecule that has a binding site for uPA; has an IC50 value
 CC of about 10 nM or less; and is not a fusion protein. Preferred molecules
 CC are uPA, (residues 1-411), single chain uPA, tcuPA (inactivated with the
 CC suicide inhibitor diisopropyl fluorophosphate), the N-terminal ATF
 CC fragment (amino acids 1-135 or 1-143) of uPA, or the growth factor domain
 CC (residues 4-43). Suitable labels include a radionuclide, a PET-imageable
 CC agent, an MRI-imageable agent, a fluorophore, a fluorochrome, a chromophore,
 CC a chromogen, a phosphorescer, a chemiluminescer or a bioluminescer. The
 CC methods are used to inhibit cell migration, cell invasion (preferably
 CC invasiveness of tumour cells), cell proliferation or angiogenesis, or to
 CC induce apoptosis, preferably in the treatment of a subject having a
 CC disease or condition associated with undesired cell migration, invasion,
 CC proliferation or angiogenesis (claimed). The protein or peptide is also
 CC useful for treating diseases or conditions including primary growth of a
 CC solid tumour, leukaemia or lymphoma, tumour invasion, metastasis,
 CC atherosclerosis, myocardial angiosclerosis, telangiectasia, corneal
 CC disease, rubeosis, neovascular glaucoma, diabetic and other retinopathy,
 CC macular degeneration, arthritis, fibrosis, wound healing with scarring
 CC and fibrosis, peptic ulcers, bone fracture, keloids, or a disorder of
 CC vasculogenesis, haematopoiesis, ovulation, menstruation, pregnancy or
 CC placentation associated with pathogenic cell invasion or with
 CC angiogenesis. The protein or peptide probe is internalised by the cells
 CC to which it binds, e.g. tumour cells, and is useful for imaging
 CC techniques in which it reduces the background signal relative to
 CC specifically bound probes. This uptake permits clearance of circulating
 CC probe so that the ratio of labelled probe inside tumour cells to the
 CC probe elsewhere in the body increases

XX Sequence 411 AA;

XX Query Match 99.0%; Score 2235; DB 4; Length 411;
 XX Best Local Similarity 97.8%; Pred. No. 5e-174;

```

: ches 402; Conservative 0; Mismatches 1; Indels 8; Gaps 1;
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCPCPKFGQCHCEIDKSKTCYEGNGHGYRG 60
Db |||||
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCPCPKFGQCHCEIDKSKTCYEQNGHGYRG 60
Db |||||
Qy 61 KASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHCNYPNPNRRRPWCYVQVGLK 120
Db |||||
Qy 61 KASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHCNYPNPNRRRPWCYVQVGLK 120
Db |||||
Qy 121 PLVQECMVHDCADGK-----LKFQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 172
Db |||||
Qy 121 PLVQECMVHDCADGKPPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 180
Db |||||
Qy 173 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 232
Db |||||
Qy 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 240
Db |||||
Qy 233 LHKDYSADTLAHNDIALKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGK 292
Db |||||
Qy 241 LHKDYSADTLAHNDIALKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGK 300
Db |||||
Qy 293 ENSTDYLYPEQLKMTVVVKLI SHRECQOPHYGSEVTTKMLCAADPOWKTDSCQDSGGPL 352
Db |||||
Qy 301 ENSTDYLYPEQLKMTVVVKLI SHRECQOPHYGSEVTTKMLCAADPOWKTDSCQDSGGPL 360
Db |||||
Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db |||||
Qy 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

```

Se ch completed: May 25, 2004, 14:52:59
 Jo time : 101.543 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

C protein - protein search, using sw model
P on: May 25, 2004, 14:53:05 ; Search time 73.4654 Seconds
      (without alignments)
      1530.046 Million cell updates/sec

T le: US-09-880-503-6
I ect score: 2257
S uence: 1 SNELHQVPSNCDLNGTCV.....VSHFLPWIRSHKTEINGLAL 403

S ring table:
Gapex 10.0 , Gapext 0.5

S rched: 1149313 seqs, 278921704 residues
T al number of hits satisfying chosen parameters: 1149313

M imum DB seq length: 0
M imum DB seq length: 2000000000

F t-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

I abase : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

F ult      Query      Match Length DB ID      Description
No.      *
-----
1 2257 100.0 403 9 US-09-880-503-6
2 2243 99.4 411 9 US-09-880-503-3
3 2243 99.4 431 12 US-10-411-026-34
4 2243 99.4 431 12 US-10-411-026-34
5 2243 99.4 431 13 US-10-076-421-2
6 2243 99.4 431 14 US-10-171-311-184
7 2243 99.4 431 14 US-10-301-822-161
8 2243 99.4 431 14 US-10-131-985-21
9 2243 99.4 431 15 US-10-395-027-414
10 2243 99.4 431 15 US-10-395-027-1275
11 2243 99.4 431 16 US-10-410-962-34
12 2243 99.4 431 16 US-10-411-049-34
13 2243 99.4 437 12 US-10-087-192-594
14 2241 99.3 431 14 US-10-247-671-149
15 2240 99.2 431 14 US-10-193-656-4

```

```

16 2233 98.9 411 15 US-10-407-821-2
17 2225 98.6 431 9 US-09-264-468B-1
18 2203 97.6 431 12 US-10-382-174-562
19 2176 96.4 445 15 US-10-360-101-266
20 1703 75.5 323 9 US-09-880-503-7
21 1626.5 72.1 433 12 US-10-087-192-591
22 1477 65.4 337 14 US-10-106-698-6266
23 1477 65.4 337 15 US-10-264-049-2927
24 1465 64.9 276 9 US-09-880-503-5
25 1464 64.9 268 15 US-10-407-821-3
26 1333 59.1 246 9 US-09-264-468B-2
27 1318 58.4 241 10 US-09-898-837A-47
28 1081.5 47.9 650 15 US-10-401-077-1
29 858.5 38.0 527 10 US-09-987-457-18
30 858.5 38.0 527 10 US-09-987-455-19
31 858.5 38.0 527 12 US-10-432-842-1
32 858.5 38.0 527 15 US-10-360-101-203
33 858.5 38.0 562 9 US-09-969-271-7
34 858.5 38.0 562 9 US-09-974-298-145
35 858.5 38.0 562 12 US-10-411-037-26
36 858.5 38.0 562 12 US-10-411-026-26
37 858.5 38.0 562 14 US-10-193-656-8
38 858.5 38.0 562 14 US-10-443-701-4
39 858.5 38.0 562 16 US-10-410-962-26
40 858.5 38.0 562 16 US-10-411-049-26
41 793 35.1 135 9 US-09-880-503-4
42 793 35.1 138 9 US-09-984-186-12
43 793 35.1 138 14 US-10-237-667-12
44 793 35.1 138 14 US-10-237-708-12
45 793 35.1 138 14 US-10-237-866-12
46 793 35.1 138 14 US-10-237-871-12
47 793 35.1 138 14 US-10-237-824-12
48 793 35.1 138 16 US-10-702-536-12
49 793 35.1 138 16 US-10-702-536-12
50 793 35.1 143 9 US-09-880-503-8
51 789.5 35.0 372 9 US-09-084-491A-3
52 789.5 35.0 372 13 US-10-102-704-3
53 784 34.7 354 10 US-09-987-457-10
54 784 34.7 354 10 US-09-987-455-11
55 784 34.7 377 10 US-09-987-455-8
56 745 33.0 343 10 US-09-987-457-14
57 745 33.0 343 10 US-09-987-455-15
58 744 33.0 339 10 US-09-987-457-12
59 744 33.0 339 10 US-09-987-455-13
60 741 32.8 331 10 US-09-987-457-11
61 741 32.8 331 10 US-09-987-455-12
62 735 32.6 343 10 US-09-987-457-15
63 735 32.6 343 10 US-09-987-455-16
64 731 32.4 335 10 US-09-987-457-13
65 731 32.4 335 10 US-09-987-455-14
66 726.5 32.2 655 14 US-10-172-712-28
67 688 30.5 308 10 US-09-987-457-16
68 688 30.5 308 10 US-09-987-455-17
69 681 30.2 615 10 US-09-858-909-2
70 681 30.2 615 14 US-10-172-712-30
71 681 30.2 615 16 US-10-449-132-2
72 652.5 28.9 560 15 US-10-391-215-6
73 651.5 28.9 560 15 US-09-912-559-3
74 651.5 28.9 560 14 US-10-172-712-32
75 651.5 28.9 560 15 US-10-391-215-5

```

ALIGNMENTS

RESULT 1

```

US-09-880-503-6
; Sequence 6, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Root
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND

```

;; TITLE OF INVENTION: TISSUE CONTRACTABILITY
;; FILE REFERENCE: 9596-331
;; CURRENT APPLICATION NUMBER: US/09/880,503
;; CURRENT FILING DATE: 2001-06-13
;; PRIOR APPLICATION NUMBER: US 60/212,847
;; PRIOR FILING DATE: 2000-06-20
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: Patent In Ver. 2.1
;; Q ID NO 6
;; LENGTH: 403
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US 9-880-503-6

Query Match 100.0%; Score 2257; DB 9; Length 403;
Best Local Similarity 100.0%; Pred. No. 1.5e-190;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60
DE 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASDTDMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDNRREPWCYVQVGLK 120
DB 61 KASDTDMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDNRREPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK-----LKFCGQKTLRPFKIIIGGEFTTIENQWFAAIYRRH 172
DB 121 PLVQECMVHDCADGKPPSEELKFCGQKTLRPFKIIIGGEFTTIENQWFAAIYRRH 180
QY 173 RGSVTVYCGSLISPCWISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLI 232
DB 181 RGSVTVYCGSLISPCWISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLI 240
QY 233 LHKDYSADTLAHDNDIALKIRSKEGRCQAQPSRTIQTICLPSMYNDPOFGTSCBITGFGK 292
DB 241 LHKDYSADTLAHDNDIALKIRSKEGRCQAQPSRTIQTICLPSMYNDPOFGTSCBITGFGK 300
QY 293 ENSTDYLYPEOLKMTVVKLISHRBCQQPHYYGSEVTTKMLCAADPQWKTDCQGDSSGGL 352
DB 301 ENSTDYLYPEOLKMTVVKLISHRBCQQPHYYGSEVTTKMLCAADPQWKTDCQGDSSGGL 360
QY 353 VCSLQGRMTLTGIYSWGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 361 VCSLQGRMTLTGIYSWGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 3
US-10-411-037-34
; Sequence 34, Application US/10411037
; Publication No. US20040043446A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zcpi, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bows, Gary
; TITLE OF INVENTION: ALPHA GALACTOSIDASE A: REMODELING AND GLYCOCONJUGATION OF ALPHA
; FILE OF INVENTION: GALACTOSIDASE A
; FILE REFERENCE: 040853-01-5082
; CURRENT APPLICATION NUMBER: US/10/411,037
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 34
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-037-34

Query Match 99.4%; Score 2243; DB 12; Length 431;
Best Local Similarity 98.1%; Pred. No. 2.8e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60

;; TITLE OF INVENTION: TISSUE CONTRACTABILITY
;; FILE REFERENCE: 9596-331
;; CURRENT APPLICATION NUMBER: US/09/880,503
;; CURRENT FILING DATE: 2001-06-13
;; PRIOR APPLICATION NUMBER: US 60/212,847
;; PRIOR FILING DATE: 2000-06-20
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: Patent In Ver. 2.1
;; Q ID NO 6
;; LENGTH: 403
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US 9-880-503-6

Query Match 100.0%; Score 2257; DB 9; Length 403;
Best Local Similarity 100.0%; Pred. No. 1.5e-190;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60
DE 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASDTDMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDNRREPWCYVQVGLK 120
DB 61 KASDTDMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDNRREPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKLPKQCCQKTLRPFKIIIGGEFTTIENQWFAAIYRRHGGSVTVV 180
DB 121 PLVQECMVHDCADGKLPKQCCQKTLRPFKIIIGGEFTTIENQWFAAIYRRHGGSVTVV 180
QY 181 CGGSLISPCWISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSAD 240
DB 181 CGGSLISPCWISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSAD 240
QY 241 TLAHNDIALKIRSKEGRCQAQPSRTIQTICLPSMYNDPOFGTSCBITGFGKENSTDYLY 300
DE 241 TLAHNDIALKIRSKEGRCQAQPSRTIQTICLPSMYNDPOFGTSCBITGFGKENSTDYLY 300
QY 301 PEOLKMTVVKLISHRBCQQPHYYGSEVTTKMLCAADPQWKTDCQGDSSGGLVCSLQGRM 360
DE 301 PEOLKMTVVKLISHRBCQQPHYYGSEVTTKMLCAADPQWKTDCQGDSSGGLVCSLQGRM 360
QY 361 TLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DE 361 TLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403

UT 2
US 9-880-503-3
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; Q ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US 9-880-503-3

Query Match 99.4%; Score 2243; DB 9; Length 411;
Best Local Similarity 98.1%; Pred. No. 2.6e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGGSEVTTKMLCAADPQWKTDSCQDGGPL 352
Db 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGGSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 353 VCSLOGRMTLTGIVSWGRGCKADKPGVYTVRSHFLPWIRSHTKENGIAL 403
Db 381 VCSLOGRMTLTGIVSWGRGCKADKPGVYTVRSHFLPWIRSHTKENGIAL 431

UT 6
US 0-171-311-184
; Tuence 184, Application US/10171311
; Slication No. US20030087270A1
; NERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; Q ID NO 184
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US 0-171-311-184

Query Match 99.4%; Score 2243; DB 14; Length 431;
Best Local Similarity 98.1%; Pred. No. 2.8e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNPKKFGQHCEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNPKKFGQHCEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQCGOKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
Db 141 PLVQECMVHDCADGKPKSPPEELKFQCGOKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200
QY 173 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLI 232
Db 201 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLI 260
QY 233 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSCETIGFGK 292
Db 261 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSCETIGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGGSEVTTKMLCAADPQWKTDSCQDGGPL 352
Db 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGGSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 353 VCSLOGRMTLTGIVSWGRGCKADKPGVYTVRSHFLPWIRSHTKENGIAL 403
Db 381 VCSLOGRMTLTGIVSWGRGCKADKPGVYTVRSHFLPWIRSHTKENGIAL 431

RESULT 7

US-10-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: WPM01-029P2RM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match 99.4%; Score 2243; DB 14; Length 431;
Best Local Similarity 98.1%; Pred. No. 2.8e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNPKKFGQHCEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNPKKFGQHCEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQCGOKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
Db 141 PLVQECMVHDCADGKPKSPPEELKFQCGOKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200
QY 173 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLI 232
Db 201 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLI 260
QY 233 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSCETIGFGK 292
Db 261 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSCETIGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGGSEVTTKMLCAADPQWKTDSCQDGGPL 352
Db 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGGSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 353 VCSLOGRMTLTGIVSWGRGCKADKPGVYTVRSHFLPWIRSHTKENGIAL 403
Db 381 VCSLOGRMTLTGIVSWGRGCKADKPGVYTVRSHFLPWIRSHTKENGIAL 431

RESULT 8

US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:

APPLICANT: Dack, Kevin N
APPLICANT: Davies, Michael J
APPLICANT: Fish, Paul V
APPLICANT: Huggins, Jonathan P
APPLICANT: McIntosh, Fraser S
APPLICANT: Occleston, Nicholas L
TITLE OF INVENTION: Composition
FILE REFERENCE: PCS 10391A
CURRENT APPLICATION NUMBER: US/10/131.985
CURRENT FILING DATE: 2002-04-25
PRIOR APPLICATION NUMBER: US/09/726.295
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: GB 9930768.8
PRIOR FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 21
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-985-21

Query Match 99.4%; Score 2243; DB 14; Length 431;
Best Local Similarity 98.1%; Pred. No. 2.8e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
61 KASDTMTGRPCLPNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
81 KASDTMTGRPCLPNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140
121 PLVQECMVHDCADGK-----LKFCGQKTLPRPKIIGGFEFTTIENQWFAAIYRRH 172
141 PLVQECMVHDCADGKSPPEELKFCGQKTLPRPKIIGGFEFTTIENQWFAAIYRRH 200
173 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 232
201 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 260
233 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGK 292
261 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGK 320
293 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVGSEVTTKMLCAADPWKTDSCQDSDGGPL 352
321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVGSEVTTKMLCAADPWKTDSCQDSDGGPL 380
353 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
381 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

ULT 9
Sequence 414, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
FILE REFERENCE: 018501-012500US

CURRENT APPLICATION NUMBER: US/10/295.027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663.733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350.666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335.394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332.464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334.393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340.376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347.211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347.349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355.250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356.714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 414
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-414

Query Match 99.4%; Score 2243; DB 15; Length 431;
Best Local Similarity 98.1%; Pred. No. 2.8e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
61 KASDTMTGRPCLPNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
81 KASDTMTGRPCLPNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140
121 PLVQECMVHDCADGK-----LKFCGQKTLPRPKIIGGFEFTTIENQWFAAIYRRH 172
141 PLVQECMVHDCADGKSPPEELKFCGQKTLPRPKIIGGFEFTTIENQWFAAIYRRH 200
173 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 232
201 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 260
233 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGK 292
261 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGK 320
293 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVGSEVTTKMLCAADPWKTDSCQDSDGGPL 352
321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVGSEVTTKMLCAADPWKTDSCQDSDGGPL 380
353 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
381 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 10
US-10-295-027-1275
Sequence 1275, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.

PPLICANT: Glynn, Richard
 PPLICANT: Hevezi, Peter A.
 PPLICANT: Mack, David H.
 PPLICANT: Murray, Richard
 PPLICANT: Watson, Susan R.
 PPLICANT: Eos Biotechnology, Inc.
 TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
 ILE REFERENCE: 018501-012500US
 URENT APPLICATION NUMBER: US/10/295,027
 URENT FILING DATE: 2002-11-13
 RIOR APPLICATION NUMBER: US 09/663,733
 RIOR FILING DATE: 2000-09-15
 RIOR APPLICATION NUMBER: US 60/350,666
 RIOR FILING DATE: 2001-11-13
 RIOR APPLICATION NUMBER: US 60/335,394
 RIOR FILING DATE: 2001-11-15
 RIOR APPLICATION NUMBER: US 60/332,464
 RIOR FILING DATE: 2001-11-21
 RIOR APPLICATION NUMBER: US 60/334,393
 RIOR FILING DATE: 2001-11-29
 RIOR APPLICATION NUMBER: US 60/340,376
 RIOR FILING DATE: 2001-12-14
 RIOR APPLICATION NUMBER: US 60/347,211
 RIOR FILING DATE: 2002-01-08
 RIOR APPLICATION NUMBER: US 60/347,349
 RIOR FILING DATE: 2002-01-10
 RIOR APPLICATION NUMBER: US 60/355,250
 RIOR FILING DATE: 2002-02-08
 RIOR APPLICATION NUMBER: US 60/356,714
 RIOR FILING DATE: 2002-02-13
 emaining Prior Application data removed - See File Wrapper or PALM.
 UMBER OF SEQ ID NOS: 1386
 OPTWARE: PatentIn Ver. 2.1
 Q ID NO 1275
 LENGTH: 431
 TYPE: PRT
 ORGANISM: Homo sapiens
 0-295-027-1275

3ry Match 99.4%; Score 2243; DB 15; Length 431;
 st Local Similarity 98.1%; Pred. No. 2.8e-189;
 tches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
 Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80

Qy 61 KASDTWGRPCLPWNSATVLOOTYHAHRS DALQLGLGKHNYCRPNDRRPPWCYVOGLK 120
 Dl 81 KASDTWGRPCLPWNSATVLOOTYHAHRS DALQLGLGKHNYCRPNDRRPPWCYVOGLK 140

Qy 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 172
 Db 141 PLVQECMVHDCADGKPPSEELKFCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 200

Qy 173 RGSVTVYCGSLSPCWVISAHCFTIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 232
 Dl 201 RGSVTVYCGSLSPCWVISAHCFTIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 260

Qy 233 LHKDYSADTLAHHNDIALLKIRSEGRCAQPSRTIQICLPSMYNDPQFTSCBITGFGK 292
 Dl 261 LHKDYSADTLAHHNDIALLKIRSEGRCAQPSRTIQICLPSMYNDPQFTSCBITGFGK 320

Qy 293 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTMMLCAADPQWKTDSCQDSSGGPL 352
 Dl 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTMMLCAADPQWKTDSCQDSSGGPL 380

Qy 353 VCSLQGRMTLGIIVSWGRGCALKDKPGYTVRVSHFLPWIRSHTKKEENGLAL 403
 Dl 381 VCSLQGRMTLGIIVSWGRGCALKDKPGYTVRVSHFLPWIRSHTKKEENGLAL 431

RESULT 11
 US-10-410-962-34
 Sequence 34, Application US/10410962
 Publication No. US20040077836A1
 GENERAL INFORMATION:
 APPLICANT: Neose Technologies, Inc.
 APPLICANT: Defrees, Shawn
 APPLICANT: Zopf, David
 APPLICANT: Bayer, Robert
 APPLICANT: Hakes, David
 APPLICANT: Chen, Xi
 APPLICANT: Bows, Caryn
 TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR: REMODELING AND
 TITLE OF INVENTION: GLYCOCONJUGATION OF G-CSF
 FILE REFERENCE: 040853-01-5054
 CURRENT APPLICATION NUMBER: US/10/410,962
 CURRENT FILING DATE: 2003-04-09
 PRIOR APPLICATION NUMBER: US 60/328,523
 PRIOR FILING DATE: 2001-10-10
 PRIOR APPLICATION NUMBER: US 60/344,692
 PRIOR FILING DATE: 2001-10-19
 PRIOR APPLICATION NUMBER: US 60/387,292
 PRIOR FILING DATE: 2002-06-07
 PRIOR APPLICATION NUMBER: US 60/391,777
 PRIOR FILING DATE: 2002-06-25
 PRIOR APPLICATION NUMBER: US 60/396,594
 PRIOR FILING DATE: 2002-07-17
 PRIOR APPLICATION NUMBER: US 60/404,249
 PRIOR FILING DATE: 2002-08-16
 PRIOR APPLICATION NUMBER: US 60/407,527
 PRIOR FILING DATE: 2002-08-28
 NUMBER OF SEQ ID NOS: 75
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 34
 LENGTH: 431
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-410-962-34

Query Match 99.4%; Score 2243; DB 16; Length 431;
 Best Local Similarity 98.1%; Pred. No. 2.8e-189;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
 Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80

Qy 61 KASDTWGRPCLPWNSATVLOOTYHAHRS DALQLGLGKHNYCRPNDRRPPWCYVOGLK 120
 Db 81 KASDTWGRPCLPWNSATVLOOTYHAHRS DALQLGLGKHNYCRPNDRRPPWCYVOGLK 140

Qy 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 172
 Db 141 PLVQECMVHDCADGKPPSEELKFCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 200

Qy 173 RGSVTVYCGSLSPCWVISAHCFTIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 232
 Db 201 RGSVTVYCGSLSPCWVISAHCFTIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 260

Qy 233 LHKDYSADTLAHHNDIALLKIRSEGRCAQPSRTIQICLPSMYNDPQFTSCBITGFGK 292
 Db 261 LHKDYSADTLAHHNDIALLKIRSEGRCAQPSRTIQICLPSMYNDPQFTSCBITGFGK 320

Qy 293 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTMMLCAADPQWKTDSCQDSSGGPL 352
 Db 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTMMLCAADPQWKTDSCQDSSGGPL 380

Qy 353 VCSLQGRMTLGIIVSWGRGCALKDKPGYTVRVSHFLPWIRSHTKKEENGLAL 403
 Db 381 VCSLQGRMTLGIIVSWGRGCALKDKPGYTVRVSHFLPWIRSHTKKEENGLAL 431

10-411-049-34
 ; Sequence 34, Application US/10411049
 ; Publication No. US20040082026A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Neose Technologies, Inc.
 ; APPLICANT: Defrees, Shawn
 ; APPLICANT: Zopf, David
 ; APPLICANT: Bayer, Robert
 ; APPLICANT: Hakes, David
 ; APPLICANT: Chen, Xi
 ; APPLICANT: Bove, Caryn
 ; TITLE OF INVENTION: INTERFERON ALPHA: REMODELING AND GLYCOCONJUGATION OF INTERFERON
 ; TITLE OF INVENTION: ALPHA
 ; FILE REFERENCE: 040853-01-5055
 ; CURRENT APPLICATION NUMBER: US/10/411,049
 ; CURRENT FILING DATE: 2003-04-09
 ; PRIOR APPLICATION NUMBER: US 60/328,523
 ; PRIOR FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 60/344,692
 ; PRIOR FILING DATE: 2001-10-19
 ; PRIOR APPLICATION NUMBER: US 60/387,292
 ; PRIOR FILING DATE: 2002-06-07
 ; PRIOR APPLICATION NUMBER: US 60/391,777
 ; PRIOR FILING DATE: 2002-06-25
 ; PRIOR APPLICATION NUMBER: US 60/396,594
 ; PRIOR FILING DATE: 2002-07-17
 ; PRIOR APPLICATION NUMBER: US 60/404,249
 ; PRIOR FILING DATE: 2002-08-16
 ; PRIOR APPLICATION NUMBER: US 60/407,527
 ; PRIOR FILING DATE: 2002-08-28
 ; NUMBER OF SEQ ID NOS: 75
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 34
 ; LENGTH: 431
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; 10-411-049-34

Query Match 99.4%; Score 2243; DB 16; Length 431;
 Best Local Similarity 98.1%; Pred. No. 2.8e-189;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 1 SNELHQVPSNCDCLNGGTCVSNKYFNSIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
 21 SNELHQVPSNCDCLNGGTCVSNKYFNSIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVQVGLK 120
 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVQVGLK 140
 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 172
 141 PLVQECMVHDCADGKPPSPPEELKFCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 200
 173 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 201 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 233 LHKDYSADTLAHHNDIALKIRKEGRCAPQSPRTIOTICLPSMYNDPQFGTSCITGFGK 292
 261 LHKDYSADTLAHHNDIALKIRKEGRCAPQSPRTIOTICLPSMYNDPQFGTSCITGFGK 320
 293 ENSTDYLYPEQLKMTVVVKLISHRECCQPHYVYSEVTTKMLCAADPQWKTDSCQDSSGGL 352
 321 ENSTDYLYPEQLKMTVVVKLISHRECCQPHYVYSEVTTKMLCAADPQWKTDSCQDSSGGL 380
 353 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 403
 381 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 431

ULT 13
 10-087-192-594

; Sequence 594, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; TITLE OF INVENTION: CANCER
 ; FILE REFERENCE: 529452000122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; CURRENT FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 594
 ; LENGTH: 437
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; 10-087-192-594

Query Match 99.4%; Score 2243; DB 12; Length 437;
 Best Local Similarity 98.1%; Pred. No. 2.8e-189;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 1 SNELHQVPSNCDCLNGGTCVSNKYFNSIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
 27 SNELHQVPSNCDCLNGGTCVSNKYFNSIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 86
 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVQVGLK 120
 87 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVQVGLK 146
 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 172
 147 PLVQECMVHDCADGKPPSPPEELKFCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 206
 173 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 207 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 266
 233 LHKDYSADTLAHHNDIALKIRKEGRCAPQSPRTIOTICLPSMYNDPQFGTSCITGFGK 292
 267 LHKDYSADTLAHHNDIALKIRKEGRCAPQSPRTIOTICLPSMYNDPQFGTSCITGFGK 326
 293 ENSTDYLYPEQLKMTVVVKLISHRECCQPHYVYSEVTTKMLCAADPQWKTDSCQDSSGGL 352
 327 ENSTDYLYPEQLKMTVVVKLISHRECCQPHYVYSEVTTKMLCAADPQWKTDSCQDSSGGL 386
 353 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 403
 387 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 437

RESULT 14
 US-10-247-671-149
 ; Sequence 149, Application US/10247671
 ; Publication No. US20030194721A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mikita, Thomas
 ; APPLICANT: Shiffman, Dov
 ; APPLICANT: Porter, Gordon, J.
 ; APPLICANT: Kaser, Matthew R.
 ; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
 ; FILE REFERENCE: PA-0050 US
 ; CURRENT APPLICATION NUMBER: US/10/247,671
 ; CURRENT FILING DATE: 2002-09-18
 ; PRIOR APPLICATION NUMBER: 60/323,784
 ; PRIOR FILING DATE: 2001-09-19
 ; NUMBER OF SEQ ID NOS: 186
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 149

```

; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US 0-247-671-149

ary Match 99.3%; Score 2241; DB 14; Length 431;
st Local Similarity 97.8%; Pred. No. 4.2e-189;
ches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSIHWNCNCPKFGQHQCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFNSIHWNCNCPKFGQHQCEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRPNDRRRPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRPNDRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 172
DB 141 PLVQECMVHDCADGKPPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 200
QY 173 RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIVVLGSRSLNSNTQGMKFEVENLI 232
DB 201 RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIVVLGSRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALLLKIRSKGRCAPSRITQITCLPSMYNDPQFGTSCBITGFGK 292
DB 261 LHKDYSADTLAHHNDIALLLKIRSKGRCAPSRITQITCLPSMYNDPQFGTSCBITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPOWKTSCQDSSGGL 352
DB 321 ENSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPOWKTSCQDSSGGL 380
QY 353 VCSLQGRMTLTCIVSWGRCALCKDKPGYVTRVSHFLPWIRSHTKXENGLAL 403
DB 381 VCSLQGRMTLTCIVSWGRCALCKDKPGYVTRVSHFLPWIRSHTKXENGLAL 431

RESULT 16
US-10-407-821-2
; Sequence 2, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
; FILE REFERENCE: UTEN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-2

Query Match 98.9%; Score 2233; DB 15; Length 411;
Best Local Similarity 97.8%; Pred. No. 2e-188;
Matches 402; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSIHWNCNCPKFGQHQCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFNSIHWNCNCPKFGQHQCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRPNDRRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRPNDRRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 172
DB 121 PLVQECMVHDCADGKPPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 180
QY 173 RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIVVLGSRSLNSNTQGMKFEVENLI 232

```

181 RGSVTVVCGSLSPCWISATHCFIDYPKKEDYIVYLGRLNSNTQGMKFEVENLI 240
233 LHKDYSADTLAHNDIALLKIRSKGRCQAPSRITQITCLPSMYNDPQGTSCITGFGK 292
241 LHKDYSADTLAHNDIALLKIRSKGRCQAPSRITQITCLPSMYNDPQGTSCITGFGK 300
293 ENSTDYLYPBQLKMTVVKLIHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDSSGGL 352
301 ENSTDYLYPBQLKMTVVKLIHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDSSGGL 360
353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKBENGLAL 403
361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKBENGLAL 411

ULT 17
US-10-282-174-562
; Sequence 562, Application US/10282174
; Publication No. US20030224380A1
; GENERAL INFORMATION:
; APPLICANT: Becker, Kenneth David
; APPLICANT: Velicelebi, Gornul
; APPLICANT: Eliios, Kathryn J.
; APPLICANT: Wang, Xin
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Bertram, Lars
; APPLICANT: Saunders, Aleister J.
; APPLICANT: Mullin, Kristina M.
; APPLICANT: Sampson, Andrew Johnson
; APPLICANT: Blacker, Deborah Lynne
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
; TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
; FILE REFERENCE: 37481-3308
; CURRENT APPLICATION NUMBER: US/10/282,174
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US 60/339,525
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 60/338,010
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/336,929
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/338,363
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/337,052
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/369,919
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 564
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 562
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 15, 58, 141, 214, 231, 274, 366
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-282-174-562

Query Match 97.6%; Score 2203; DB 12; Length 431;
Best Local Similarity 96.6%; Pred. No. 9.4e-186;
Matches 397; Conservative 0; Mismatches 6; Indels 8; Gaps 1;

1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGHCIDKSKTCYEGNGHFYRG 60
21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGHCIDKSKTCYEGNGHFYRG 80
61 KASDTMTGRPCLPNSATVLTQYTHAHRSDALQGLGKHNYCRPNRRRPPWCYVQVGLK 120
81 KASDTMTGRPCLPNSATVLTQYTHAHRSDALQGLGKHNYCRPNRRRPPWCYVQVGLK 140
121 PLVQECMVHDCADGK-----LKFCQCKTIRPRFKIIGGFTTIENQWFAALYRRH 172
141 XLVQECMVHDCADGKPPPEELKFCQCKTIRPRFKIIGGFTTIENQWFAALYRRH 200

181 RGSVTVVCGSLSPCWISATHCFIDYPKKEDYIVYLGRLNSNTQGMKFEVENLI 240
233 LHKDYSADTLAHNDIALLKIRSKGRCQAPSRITQITCLPSMYNDPQGTSCITGFGK 292
241 LHKDYSADTLAHNDIALLKIRSKGRCQAPSRITQITCLPSMYNDPQGTSCITGFGK 300
293 ENSTDYLYPBQLKMTVVKLIHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDSSGGL 352
301 ENSTDYLYPBQLKMTVVKLIHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDSSGGL 360
353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKBENGLAL 403
361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKBENGLAL 411

ULT 17
US-264-468B-1
; Sequence 1, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.P1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
; OTHER INFORMATION: Leader sequence
; NAME/KEY: VARIANT
; LOCATION: (279)...(279)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (302)...(302)
; OTHER INFORMATION: Xaa = any amino acid
US-264-468B-1

Query Match 98.6%; Score 2225; DB 9; Length 431;
Best Local Similarity 97.6%; Pred. No. 1.1e-187;
Matches 403; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGHCIDKSKTCYEGNGHFYRG 60
21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGHCIDKSKTCYEGNGHFYRG 80
61 KASDTMTGRPCLPNSATVLTQYTHAHRSDALQGLGKHNYCRPNRRRPPWCYVQVGLK 120
81 KASDTMTGRPCLPNSATVLTQYTHAHRSDALQGLGKHNYCRPNRRRPPWCYVQVGLK 140
121 PLVQECMVHDCADGK-----LKFCQCKTIRPRFKIIGGFTTIENQWFAALYRRH 172
141 PLVQECMVHDCADGKPPPEELKFCQCKTIRPRFKIIGGFTTIENQWFAALYRRH 200
173 RGSVTVVCGSLSPCWISATHCFIDYPKKEDYIVYLGRLNSNTQGMKFEVENLI 232
201 RGSVTVVCGSLSPCWISATHCFIDYPKKEDYIVYLGRLNSNTQGMKFEVENLI 260

QY 173 RGGSVTVCGGSLISPCWISATHCFTDYPKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 DE 201 RGGSVTVCGGSLISPCWISATHCFTDYPKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
 DE 261 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
 QY 293 ENSTDYLYPEOLKMTVVKLISHRECOQPHYGSEVTTMMLCAADPQWKTDSCQDGGPL 352
 DE 321 ENSTDYLYPEOLKMTVVKLISHRECOQPHYGSEVTTMMLCAADPQWKTDSCQDGGPL 380
 QY 353 VCSLQGRMTLGTIVSWGRGCKADKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
 DE 381 VCSLQGRMTLGTIVSWGRGCKADKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RE UT 19
 US 0-360-101-266
 ; Sequence 266, Application US/10360101
 ; Publication No. US20040009550A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Moll, Gert N.
 ; PPLICANT: Leenhouts, Cornelis J.
 ; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
 ; FILE REFERENCE: 2183-5673
 ; CURRENT APPLICATION NUMBER: US/10/360.101
 ; PRIOR FILING DATE: 2003-02-07
 ; PRIOR APPLICATION NUMBER: EP 02077060.8
 ; PRIOR FILING DATE: 2002-05-24
 ; NUMBER OF SEQ ID NOS: 309
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 266
 ; LENGTH: 445
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: sequence of urokinase
 US 0-360-101-266

QY 173 RGGSVTVCGGSLISPCWISATHCFTDYPKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 DE 201 RGGSVTVCGGSLISPCWISATHCFTDYPKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
 DE 261 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
 QY 293 ENSTDYLYPEOLKMTVVKLISHRECOQPHYGSEVTTMMLCAADPQWKTDSCQDGGPL 352
 DE 321 ENSTDYLYPEOLKMTVVKLISHRECOQPHYGSEVTTMMLCAADPQWKTDSCQDGGPL 380
 QY 353 VCSLQGRMTLGTIVSWGRGCKADKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
 DE 381 VCSLQGRMTLGTIVSWGRGCKADKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RE UT 19
 US 0-360-101-266
 ; Sequence 266, Application US/10360101
 ; Publication No. US20040009550A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Moll, Gert N.
 ; PPLICANT: Leenhouts, Cornelis J.
 ; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
 ; FILE REFERENCE: 2183-5673
 ; CURRENT APPLICATION NUMBER: US/10/360.101
 ; PRIOR FILING DATE: 2003-02-07
 ; PRIOR APPLICATION NUMBER: EP 02077060.8
 ; PRIOR FILING DATE: 2002-05-24
 ; NUMBER OF SEQ ID NOS: 309
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 266
 ; LENGTH: 445
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: sequence of urokinase
 US 0-360-101-266

QY 399 NGLAL 403
 DB 441 NGLAL 445

RESULT 20
 US-09-880-503-7
 ; Sequence 7, Application US/09880503
 ; Patent No. US20020131964A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHINESE, Douglas B
 ; APPLICANT: HIGAZI, Add Al-Roof
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
 ; TITLE OF INVENTION: TISSUE CONTRACTILITY
 ; FILE REFERENCE: 9596-331
 ; CURRENT APPLICATION NUMBER: US/09/880,503
 ; CURRENT FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/212,847
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 7
 ; LENGTH: 323
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-880-503-7

Query Match 75.5%; Score 1703; DB 9; Length 323;
 Best Local Similarity 79.2%; Pred. No. 8.2e-142;
 Matches 319; Conservative 2; Mismatches 2; Indels 80; Gaps 3;

QY 1 SNELHQPVSNCDCCLNGGTCVSNKYFNSNIHWCNCPKFGQHCEDKSKTCYEGNGHYFG 60
 DB 1 SNELHQPVSNCDCCLNGGTCVSNKYFNSNIHWCNCPKFGQHCEDKSKTCYEGNGHYFG 48
 QY 61 KASTDTMGRPCLPWNSATVLTQTYFTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQ 120
 DB 49 -----PSSP-----PE----- 54

QY 121 PLVQECMVHDCADGKLPKFCQCGOKTLRPRFKIIGBEFTTIENQPFALYRHRGGSVTV 180
 DB 55 -----ELKFCQCGOKTLRPRFKIIGBEFTTIENQPFALYRHRGGSVTV 100

QY 181 CGSLISPCWISATHCFTDYPKEDYIVYLGSRSLNSNTQGMKFEVENLIHKDYSAD 240
 DB 101 CGSLISPCWISATHCFTDYPKEDYIVYLGSRSLNSNTQGMKFEVENLIHKDYSAD 160

QY 241 TLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGKENSTDYLY 300
 DB 161 TLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGKENSTDYLY 220

QY 301 PSOLKMTVVKLISHRECOQPHYGSEVTTMMLCAADPQWKTDSCQDGGPLVCSLQGRM 360
 DB 221 PSOLKMTVVKLISHRECOQPHYGSEVTTMMLCAADPQWKTDSCQDGGPLVCSLQGRM 280

QY 361 TLTGIVSWGRGCKADKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
 DB 281 TLTGIVSWGRGCKADKPGVYTRVSHFLPWIRSHTKKEENGLAL 323

RESULT 21
 US-10-087-192-591
 ; Sequence 591, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; TITLE OF INVENTION: CANCER
 ; FILE REFERENCE: 529452000122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; CURRENT FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377

PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 2059
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 591
LENGTH: 433
TYPE: PRT
ORGANISM: Mus musculus
10-087-192-591

Query Match 72.1%; Score 1626.5; DB 12; Length 433;
est Local Similarity 70.2%; Pred. No. 6.6e-135;
atches 283; Conservative 50; Mismatches 61; Indels 9; Gaps 2;

9 SNCDLNGGTCVSNKYFNSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFGYRG 68
30 SNCCQNGGVCVSYKIFSRIRRCSPRKFQEGHCEIDASKTCYHGNGDSYRGKANTDTKG 89
69 RPCLPWNASATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRNRRPWCYVQVGLKPLVQECMV 128
90 RPCLAWAPALQKPYNAHRPDAISLGLGKHNYCRNPDNRNRRPWCYVQVGLRQFVQECMV 149
129 HDCADGK-----FQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRHGS-VTY 179
150 HDGSLSKPSSVDQCGQKALRPRFKIIGGEFTTIENQWFAAIYRRHGS-VTY 209
180 VCGSLISPCWISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKPEVENLILHKDYSA 239
210 KCGSLISPCWASAAHCFIOLPKENYVYVLSKSSSYNGEMKFEVQLILHYYRE 269
240 DTLAHDNDIALKIRSGRCAQPSRTIOTCLPSMNDPQPTSCETITGFGKENTDYL 299
270 DSLAHDNDIALKIRSGRCAQPSRTIOTCLPSMNDPQPTSCETITGFGKESDYL 329
300 YPELQKMTVVKLISHRSCQPHYYSVTTKMLCAADPQWKTSCOGDSGGLVCSLQGR 359
330 YPNLKMVSVKLSHEQCMQPHYYSVTTKMLCAADPQWKTSCOGDSGGLVCSLQGR 389
360 MTLTGIVSGRGKALDKPKGYTRVSHFLPWIRSHTKENGLA 402
390 PTLGIVSGRGKAEKKNKPGYTRVSHFLPWIRSHTKENGLA 432

ULT 22
10-106-698-6266
Sequence 6266, Application US/10106698
Publication No. US2003010960A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
FILE REFERENCE: PA00591
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: Patent in Ver. 3.0
SEQ ID NO 6266
LENGTH: 337
TYPE: PRT
ORGANISM: Homo sapiens
10-106-698-6266

Query Match 65.4%; Score 1477; DB 14; Length 337;
est Local Similarity 92.4%; Pred. No. 7.4e-122;
atches 267; Conservative 3; Mismatches 5; Indels 14; Gaps 2;

1 SNELHQVPSNCDLNGGTCVSNKYFNSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFGYRG 60

Db 27 SNELHQVPSNCDLNGGTCVSNKYFNSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFGYRG 86
QY 61 KASDTMTGRPCLPWNASATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
Db 87 KASDTMTGRPCLPWNASATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 146
QY 121 PLVQECMVHDCADGK-----LKFCQGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
Db 147 PLVQECMVHDCADGKPKSSPEELKFCQGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 206
QY 173 RGGSVTYVCGGSLISPCWISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKPEVENLI 232
Db 207 RGGSVTYVCGGSLISPCWISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKPEVENLI 266
QY 233 LHKDYSADTLAHDNDIALKIRSGRCAQ-----PSRTIOTCLPSM 275
Db 267 LHKDYSADTLAHDNDIALKIRSGRCAQHPGLYRFSACPRCITIPSL 315

RESULT 23
US-10-264-049-2927
Sequence 2927, Application US/10264049
Publication No. US20040005579A1
GENERAL INFORMATION:
APPLICANT: Birse et al.
FILE REFERENCE: PA133P1
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: Patent in Ver. 3.1
SEQ ID NO 2927
LENGTH: 337
TYPE: PRT
ORGANISM: Homo sapiens
US-10-264-049-2927

Query Match 65.4%; Score 1477; DB 15; Length 337;
est Local Similarity 92.4%; Pred. No. 7.4e-122;
Matches 267; Conservative 3; Mismatches 5; Indels 14; Gaps 2;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFNSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFGYRG 60
Db 27 SNELHQVPSNCDLNGGTCVSNKYFNSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFGYRG 86
QY 61 KASDTMTGRPCLPWNASATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
Db 87 KASDTMTGRPCLPWNASATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 146
QY 121 PLVQECMVHDCADGK-----LKFCQGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
Db 147 PLVQECMVHDCADGKPKSSPEELKFCQGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 206
QY 173 RGGSVTYVCGGSLISPCWISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKPEVENLI 232
Db 207 RGGSVTYVCGGSLISPCWISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKPEVENLI 266
QY 233 LHKDYSADTLAHDNDIALKIRSGRCAQ-----PSRTIOTCLPSM 275
Db 267 LHKDYSADTLAHDNDIALKIRSGRCAQHPGLYRFSACPRCITIPSL 315

RESULT 24
US-09-880-503-5
Sequence 5, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B

```

; APPLICANT: HIGAZI, Abd Al-Boof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; RIOR APPLICATION NUMBER: US 60/212,847
; RIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; Q ID NO 5
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US 9-880-503-5

ary Match 64.9%; Score 1465; DB 9; Length 276;
st Local Similarity 99.6%; Pred. No. 5.6e-121;
tches 268; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 135 KLFQCGQKTLRPRFKIIGGFTTIENQWFAAIYRHRGGSVTVVCGSLISPCWISA 194
Db 8 ELKFCGQKTLRPRFKIIGGFTTIENQWFAAIYRHRGGSVTVVCGSLISPCWISA 67

Qy 195 THCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIR 254
De 68 THCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIR 127

Qy 255 SKEGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMTVVKLISH 314
De 128 SKEGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMTVVKLISH 187

Qy 315 RECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGCAL 374
De 188 RECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGCAL 247

Qy 375 KDKPGVYTRVSHFLPWIRSHKTEENGLAL 403
De 248 KDKPGVYTRVSHFLPWIRSHKTEENGLAL 276

RE UT 25
US 0-407-821-3
; Sequence 3, Application US/10407821
; Patent No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; FILE REFERENCE: UTSN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; RIOR FILING DATE: 2003-04-04
; RIOR APPLICATION NUMBER: 60/414,202
; RIOR FILING DATE: 2002-09-27
; RIOR APPLICATION NUMBER: 60/370,466
; RIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in Ver. 2.1
; Q ID NO 3
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US 0-407-821-3

ary Match 64.9%; Score 1464; DB 15; Length 268;
st Local Similarity 100.0%; Pred. No. 7.9e-121;
tches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 LKFCGQKTLRPRFKIIGGFTTIENQWFAAIYRHRGGSVTVVCGSLISPCWISAT 195
De 1 LKFCGQKTLRPRFKIIGGFTTIENQWFAAIYRHRGGSVTVVCGSLISPCWISAT 60

```

```

Qy 196 HCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRS 255
Db 61 HCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRS 120

Qy 256 KEGRCQAQPSRTIQTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMTVVKLISHR 315
Db 121 KEGRCQAQPSRTIQTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMTVVKLISHR 180

Qy 316 EQQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGCAL 375
Db 181 EQQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGCAL 240

Qy 376 KDPGVYTRVSHFLPWIRSHKTEENGLAL 403
Db 241 KDPGVYTRVSHFLPWIRSHKTEENGLAL 268

RESULT 26
US-09-264-468B-2
; Sequence 2, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.P1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-264-468B-2

Query Match 59.1%; Score 1333; DB 9; Length 246;
Best Local Similarity 99.2%; Pred. No. 2.5e-109;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 151 IIGGFTTIENQWFAAIYRHRGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVY 210
Db 1 IIGGFTTIENQWFAAIYRHRGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVY 60

Qy 211 LGRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQT 270
Db 61 LGRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQT 120

Qy 271 CLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTK 330
Db 121 ALPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTK 180

Qy 331 MLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPW 390
Db 181 MLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPW 240

Qy 391 IRSHTK 396
Db 241 IRSHTK 246

RESULT 27
US-09-898-837A-47
; Sequence 47, Application US/09898837A

```



```
Publication No. US20030077697A1
GENERAL INFORMATION:
APPLICANT: Quinn, Kerry E.
APPLICANT: Spytak, Kimberly A.
APPLICANT: Majumder, Kumud
APPLICANT: Vernet, Corine
APPLICANT: Herrmann, John L.
APPLICANT: Burgess, Catherine
APPLICANT: Fernandes, Elma
APPLICANT: Taupier Jr., Raymond
APPLICANT: Rastelli, Luca
APPLICANT: Curagen Corporation
APPLICANT: Gerlach, Valerie L
APPLICANT: MacDougall, John R
TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND
FILE REFERENCE: NUCLEIC ACIDS ENCODING THE SAME
CURRENT APPLICATION NUMBER: US/09/898,837A
CURRENT FILING DATE: 2001-07-03
PRIORITY APPLICATION NUMBER: U.S.S.N. 60/165,986
PRIORITY FILING DATE: 1999-11-17
PRIORITY APPLICATION NUMBER: U.S.S.N. 60/194,839
PRIORITY FILING DATE: 2000-04-05
PRIORITY APPLICATION NUMBER: U.S.S.N. 60/195,637
PRIORITY FILING DATE: 2000-04-07
PRIORITY APPLICATION NUMBER: U.S.S.N. 60/197,080
PRIORITY FILING DATE: 2000-04-13
PRIORITY APPLICATION NUMBER: U.S.S.N. 60/232,677
PRIORITY FILING DATE: 2000-09-15
PRIORITY APPLICATION NUMBER: U.S.S.N. 60/181,347
PRIORITY FILING DATE: 2000-02-09
PRIORITY APPLICATION NUMBER: U.S.S.N. 60/194,195
PRIORITY FILING DATE: 2000-04-03
PRIORITY APPLICATION NUMBER: U.S.S.N. 60/215,906
PRIORITY FILING DATE: 2000-07-03
PRIORITY APPLICATION NUMBER: U.S.S.N. 09/715,427
PRIORITY FILING DATE: 2000-11-16
NUMBER OF SEQ ID NOS: 53
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 47
LENGTH: 241
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-837A-47

Query Match 58.4%; Score 1318; DB 10; Length 241;
Best Local Similarity 99.6%; Pred. No. 5.2e-108;
Matches 240; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

151 IIGGETTTIENQWPFAAIYRRHGGSVTVVCGSLSPCWVISATHCFIDYPKKEDIYVY 210
|||||
1 IIGGETTTIENQWPFAAIYRRHGGSVTVVCGSLSPCWVISATHCFIDYPKKEDIYVY 60

211 LGSRLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALLKIRSGRCAQPSRTIOTI 270
|||||
61 LGSRLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALLKIRSGRCAQPSRTIOTI 120

271 CLPSMTNDPQFGTSCITGFKENSTDYLYPEQLKMTVVKLISHRECQPHYGVSEVTK 330
|||||
121 CLPSMTNDPQFGTSCITGFKENSTDYLYPEQLKMTVVKLISHRECQPHYGVSEVTK 180

331 MLCADDPQWKTDSCQDGGPLVCSLOGRMTLTGIVSGRGKALDKPGVYTRVSHFLPW 390
|||||
181 MLCADDPQWKTDSCQDGGPLVCSLOGRMTLTGIVSGRGKALDKPGVYTRVSHFLPW 240

391 I 391
241 I 241
```

```
Publication No. US20040002137A1
GENERAL INFORMATION:
APPLICANT: Hung, Paul Porwen
APPLICANT: Wu, Bryan T. H.
TITLE OF INVENTION: HUMAN TISSUE UROKINASE TYPE PLASMINOGEN
FILE REFERENCE: 12133-006001
CURRENT APPLICATION NUMBER: US/10/401,077
CURRENT FILING DATE: 2003-03-27
PRIORITY APPLICATION NUMBER: US 60/371,013
PRIORITY FILING DATE: 2002-04-09
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 650
TYPE: PRT
ORGANISM: Homo sapiens
US-10-401-077-1

Query Match 47.9%; Score 1081.5; DB 15; Length 650;
Best Local Similarity 39.2%; Pred. No. 1.3e-86;
Matches 226; Conservative 49; Mismatches 116; Indels 185; Gaps 10;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYPFSNIHWCNPKKFGGQCEIDKSKTCYEGNGHY 58
DB 77 QCHSVPVKSCSEPRCFNGGTCQQALYFSDP-VQCQEPGFAGKCEIDTRATCTYEGNGHY 135
QY 59 RGKASTDTMGRPCLPWNSATVLOOTYHAHRSDALQLGLGHKHYCRNPDNRRRPMWCYVQVG 118
DB 136 RGKASTDTMGRPCLPWNSATVLOOTYHAHRSDALQLGLGHKHYCRNPDNRRRPMWCYVQVG 195
QY 119 LKPLVQECMVHDCADG----- 134
DB 196 LKPLVQECMVHDCSEGNSDCYEQGISYRGTSWSTAESGAECTWNSSALAKQPYSGRRPD 255
QY 135 ----- 134
DB 256 AIRLGLGNHYCRNPDNRDPSKPCVFKAGKYSSEFCSTPACSEGNSDCYFNGSAYRGTH 315
QY 135 -----KLK 137
DB 316 SLTESGASCLPWNMILIKVYTAQNPDAALGLGHKHYCRNPDGDAKPCWCHVKNRRLT 375
QY 138 FQ-----CG-QKTLRPRFKIIGGEFTTIENQWPFAAIYRRH-RGGSVTVVCGGSLI 186
DB 376 WEYCDVPSCSTGLRQYSQPFRIKGLPADIASHPWQAAIFAKHRESGPERFLCGGILI 435
QY 187 SPQWVIAETHCFIDYPKKEDIYVILGSRNSNTQGMKFEVENLILHKDYSADTLAHHN 246
DB 436 SSCMILSAAHCFQERFPFPHLTIVLGRTYRVVPEEEOKEFEVKYIVHKEFDDDT--YDN 493
QY 247 DIALLKIRSGRCAQPSRTIOTICLPSPMYNDPQFGTSCITGFKENSTDYLYPEOLKM 306
DB 494 DIALLQLKSDSSRCAQSSVVTVLPADLQLPDWTCELSGVGKHEALSPFYSERLKE 553
QY 307 TVVKLISHRCQPHYGVSEVTKMLCAAD-----POWKT-DSQCGDGGPLVCSLOGRM 360
DB 554 AHVELYSESRCTSCHLLNRTVTDNMLCAGTRSGGPOANLHDACQDGGSGGFLVCLNDGRM 613
QY 361 TLTCIVSGRGKALDKPGVYTRVSHFLPWIRSHTK 396
DB 614 TLVGIIISGLGCGQKQDVPGYTKVTNYLDWIRDMR 649

RESULT 29
US-09-987-457-18
Sequence 18, Application US/09987457
Publication No. US20030013150A1
GENERAL INFORMATION:
APPLICANT: Manosroi, Aranya
APPLICANT: Manosroi, Jiradej
APPLICANT: Tayapiwatana, Chatchai
APPLICANT: Goetz, Friedrich
```

```

; PPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,457
; PRIORITY FILING DATE: 2001-11-14
; RIOR APPLICATION NUMBER: 60/268,573
; RIOR FILING DATE: 2001-02-15
; RIOR APPLICATION NUMBER: GB 00 27 782.2
; RIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; ORGANISM: Homo sapiens (tpe)
; US-987-457-18

Query Match      38.0%; Score 858.5; DB 10; Length 527;
Best Local Similarity 37.5%; Pred. No. 4.5e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

QY 3 ELHVP-SNCD---CLNGGTCVSNKYSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHGY 58
DB 42 QCHSVPVKSCSEPCFNGGTCQALYFSDP-VQCPGEGFAGKCEIDTRATCYEDQGISY 100
QY 59 RGASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRKCYVQVG 118
DB 101 RGTWSTAESGAECTNWNSSALAQKPYSGRRPDARLGLGNHNYCRNPDNRDSKPCWYVFK 160
QY 119 LKPLVQECMVHDCADG----- 134
DB 161 GKYSSEFCSTPACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKYVTAQNP 220
QY 135 -----KLKFO----- 155
DB 221 AQALGLGKHNYCRNPDGDAPKFWCHLVKNRRLTWEYCDVPSCSTCGLRQYSPQFRKGG 280
QY 156 FTTTENQWFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVLGRS 214
DB 281 PADTASHPWQAIAFAKRRSPGERFLCGGILISSCWLSAAHCFQERFPPHLLTVILGRT 340
QY 215 RLNSNTOEMKFVENILHKDYSADTLAHNDIALKIRSKGRCACQPSRTIOTICLPS 274
DB 341 YRVVPGEEQKFEVEKYIVHKEFDDT--YNDNIALQLKSDSRCAQESSVVRTVCLPP 398
QY 275 MYNDPQFTSCETITGFGKENSTDVLYPEOLKMTVVKLISHRECQOPHYGSEVTTKMLCA 334
DB 399 ADLQLPDWTCELSGKHEALSPFYSERLKEAHRVLYPSSRCTSQHLLNRTVTDNMLCA 458
QY 335 AD-----PWKT-DSCQSDGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFL 388
DB 459 GDRSGGQPANLHDACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDPVPGVYTKVTNYL 518
QY 389 PWIRSHTK 396
DB 519 DWIRDNR 526

RESULT 31
US-10-432-842-1
; Sequence 1, Application US/10432842
; Publication No. US20040071707A1
; GENERAL INFORMATION:
; APPLICANT: Veronica A. CARROLL
; APPLICANT: Adrian L. HARRIS
; APPLICANT: Roy BICKNELL
; APPLICANT: Pat PRICE
; TITLE OF INVENTION: MODULATION OF CELL GROWTH
; FILE REFERENCE: 117-450 / N.79507A SER
; CURRENT APPLICATION NUMBER: US/10/432,842
; CURRENT FILING DATE: 2003-09-27
; PRIOR APPLICATION NUMBER: PCT/GB01/05244
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: GB 0029001.5
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: MS Word
; SEQ ID NO 1
```

```

; PPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,457
; PRIORITY FILING DATE: 2001-11-14
; RIOR APPLICATION NUMBER: 60/268,573
; RIOR FILING DATE: 2001-02-15
; RIOR APPLICATION NUMBER: GB 00 27 782.2
; RIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; ORGANISM: Homo sapiens (tpe)
; US-987-457-18

Query Match      38.0%; Score 858.5; DB 10; Length 527;
Best Local Similarity 37.5%; Pred. No. 4.5e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

QY 3 ELHVP-SNCD---CLNGGTCVSNKYSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHGY 58
DB 42 QCHSVPVKSCSEPCFNGGTCQALYFSDP-VQCPGEGFAGKCEIDTRATCYEDQGISY 100
QY 59 RGASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRKCYVQVG 118
DB 101 RGTWSTAESGAECTNWNSSALAQKPYSGRRPDARLGLGNHNYCRNPDNRDSKPCWYVFK 160
QY 119 LKPLVQECMVHDCADG----- 134
DB 161 GKYSSEFCSTPACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKYVTAQNP 220
QY 135 -----KLKFO----- 155
DB 221 AQALGLGKHNYCRNPDGDAPKFWCHLVKNRRLTWEYCDVPSCSTCGLRQYSPQFRKGG 280
QY 156 FTTTENQWFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVLGRS 214
DB 281 PADTASHPWQAIAFAKRRSPGERFLCGGILISSCWLSAAHCFQERFPPHLLTVILGRT 340
QY 215 RLNSNTOEMKFVENILHKDYSADTLAHNDIALKIRSKGRCACQPSRTIOTICLPS 274
DB 341 YRVVPGEEQKFEVEKYIVHKEFDDT--YNDNIALQLKSDSRCAQESSVVRTVCLPP 398
QY 275 MYNDPQFTSCETITGFGKENSTDVLYPEOLKMTVVKLISHRECQOPHYGSEVTTKMLCA 334
DB 399 ADLQLPDWTCELSGKHEALSPFYSERLKEAHRVLYPSSRCTSQHLLNRTVTDNMLCA 458
QY 335 AD-----PWKT-DSCQSDGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFL 388
DB 459 GDRSGGQPANLHDACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDPVPGVYTKVTNYL 518
QY 389 PWIRSHTK 396
DB 519 DWIRDNR 526

RESULT 31
US-10-432-842-1
; Sequence 1, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; PPLICANT: Aranya Manosroi
; PPLICANT: Jiradej Manosroi
; PPLICANT: Chatchai Tayapiwatana
; PPLICANT: Friedrich Goetz
; PPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
```

```

; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; 10-432-842-1
;
; Query Match
; est Local Similarity 38.0%; Score 858.5; DB 12; Length 527;
; atches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
;
; 3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKKFGQGHCEIDKSKTCYEGNGHGY 58
; 42 QCHSVPVKSCSEPRCFNGGTCQOALYFSDP--VQCPEGFAGKCEIDTRATCYEDQGISY 100
;
; 59 RGKASTDTMGRPCLPWNATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVYQVG 118
; 101 RGTWSTAESGAECTNWNSSALAQKPYSGRRPDALRLGLGNHNYCRNPDNRDRSKPWCYVKA 160
;
; 119 LKPLVQECMVHDCADG----- 134
; 161 GKYSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKYVTAQNP 220
; 135 -----KLKFO----- 155
; 221 AALGLGKHNYCRNPDGAKPCHLVKNRRLTWECYDVPSCSTCGLRQYSQPFRIKGG 280
; 156 FTTTENOPFAAIYRRH--RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRS 214
; 281 FADTASHPWQAIAI FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHLLTVILGRT 340
; 215 RLNSNTQGMKFVENILHKDYSADTLAHNDIALKIRSGRCACQAPRTIQTICLPS 274
; 341 YRVVPGSEEQKFEVEKYIVHKEFDDT--YNDNIALQLKSDSRCAQESSVVRTVCLPP 398
; 275 MYNDPQFCTSCEITGFGKENS TDLYPEQLKMTVVKLIHRECCQPHYVSGEVTVMCLCA 334
; 399 ADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRCTSQHLLNRTVTDNMLCA 458
; 335 AD-----POWKT--DSCQDGGPLVCSLQGRMTLTGIVSWGRCALDKPGVTRVSHFL 388
; 459 GDTSGGGPQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLGCGQKDVPGVYTKVTNYL 518
;
; 389 PWIRSHTK 396
; 519 DWIRDNMR 526
;
; ULT 32
; 10-360-101-203
; sequence 203, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360.101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 203
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: sequence of alteplase
; 10-360-101-203
;
; Query Match
; est Local Similarity 38.0%; Score 858.5; DB 15; Length 527;
; atches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

```

```

; 3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKKFGQGHCEIDKSKTCYEGNGHGY 58
; 42 QCHSVPVKSCSEPRCFNGGTCQOALYFSDP--VQCPEGFAGKCEIDTRATCYEDQGISY 100
;
; 59 RGKASTDTMGRPCLPWNATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVYQVG 118
; 101 RGTWSTAESGAECTNWNSSALAQKPYSGRRPDALRLGLGNHNYCRNPDNRDRSKPWCYVKA 160
;
; 119 LKPLVQECMVHDCADG----- 134
; 161 GKYSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKYVTAQNP 220
; 135 -----KLKFO----- 155
; 221 AALGLGKHNYCRNPDGAKPCHLVKNRRLTWECYDVPSCSTCGLRQYSQPFRIKGG 280
; 156 FTTTENOPFAAIYRRH--RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRS 214
; 281 FADTASHPWQAIAI FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHLLTVILGRT 340
; 215 RLNSNTQGMKFVENILHKDYSADTLAHNDIALKIRSGRCACQAPRTIQTICLPS 274
; 341 YRVVPGSEEQKFEVEKYIVHKEFDDT--YNDNIALQLKSDSRCAQESSVVRTVCLPP 398
; 275 MYNDPQFCTSCEITGFGKENS TDLYPEQLKMTVVKLIHRECCQPHYVSGEVTVMCLCA 334
; 399 ADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRCTSQHLLNRTVTDNMLCA 458
; 335 AD-----POWKT--DSCQDGGPLVCSLQGRMTLTGIVSWGRCALDKPGVTRVSHFL 388
; 459 GDTSGGGPQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLGCGQKDVPGVYTKVTNYL 518
;
; 389 PWIRSHTK 396
; 519 DWIRDNMR 526
;
; RESULT 33
; US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-969-271-7
;
; Query Match
; est Local Similarity 38.0%; Score 858.5; DB 9; Length 562;
; atches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
;
; 3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKKFGQGHCEIDKSKTCYEGNGHGY 58
; 77 QCHSVPVKSCSEPRCFNGGTCQOALYFSDP--VQCPEGFAGKCEIDTRATCYEDQGISY 135
;
; 59 RGKASTDTMGRPCLPWNATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVYQVG 118
; 136 RGTWSTAESGAECTNWNSSALAQKPYSGRRPDALRLGLGNHNYCRNPDNRDRSKPWCYVKA 195
;
; 119 LKPLVQECMVHDCADG----- 134
; 196 GKYSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKYVTAQNP 255

```

QY 135 -----KLKQ-----CG-QKTLRPRFKIIGGE 155
Db 256 AQAALGLGKHNCRNPDGDAKFWCHVLKNRRLTWECYCDVPSCTGGLRQYSQPQFRIKGL 315
QY 156 FTTIENQWFAATYRRH-RGGSVTVYCGSLISPCWVIGATHCFIDYPKKEDYIVVLGRS 214
Db 316 FADIASHPQAAIFAKHRRSPGERFLCGGILLSSCWILSAAHCFQFRPPHLLTILGRT 375
QY 215 RLNSNTGEMKFEVENLILHKDYSADTLAHNDIALLKIRSKRCAQPSRTIOTICLPS 274
Db 376 YRVVPGEEQKFEVEKIVHKEFDDT--YNDIALLLQKSSRCAQESSVVRTVCLPP 433
QY 275 MYNDPQGTSCBITGFKENSTDYLPBOLKMTVVKLISHRECOQPHYGVSEVTTMCLCA 334
Db 434 ADLQLPDWTCELSGYKHEALSPFYSERLKEARVLYPSSRCTSQHLLNRTVTDNMLCA 493
QY 335 AD-----POWKT-DSCOGDSGGPLVCSLOGRMILTGTIVSNGRCALKDKFQVTVRVSHFL 388
Db 494 GDTSGGFPQANLHDACQDSGGPLVCLNDGRMTLVGLISWGLGCGQKDVGVTVKVTNYL 553
QY 389 PWIRSHTK 396
Db 554 DWIRDNR 561
RESULT 35
US-10-411-037-26
; Sequence 26, Application US/10411037
; Publication No. US2004004346A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bove, Caryn
; TITLE OF INVENTION: ALPHA GALACTOSIDASE A: REMODELING AND GLYCOCONJUGATION OF ALPHA
; FILE OF INVENTION: GALACTOSIDASE A
; FILE REFERENCE: 040853-01-5082
; CURRENT APPLICATION NUMBER: US/10/411,037
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-037-26
Query Match 38.0%; Score 858.5; DB 9; Length 562;
Best Local Similarity 37.5%; Pred. No. 4.9e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGHCIDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSPRCFNGGTCQALYFSDF-VCQCPGFGAGKCCBIDTRATCYEDQGISY 135
QY 59 RKGASTDTMGRCLPWNSTVLTQTYHAHRSDALQGLGKHNCRNPNRPRKCYQVQ 118
Db 136 RGTWSTAESGAECTNWNSSALAKPKYSGRRPDAIRLGLGNHNYCRNPDSDSKPVCYFKA 195
QY 119 LKPLVQECMVHDCADG----- 134
Db 196 GKYSSFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLRWNSMILIGKVTYQNP 255
QY 135 -----KLKQ-----CG-QKTLRPRFKIIGGE 155
Db 256 AQAALGLGKHNCRNPDGDAKFWCHVLKNRRLTWECYCDVPSCTGGLRQYSQPQFRIKGL 315
QY 156 FTTIENQWFAATYRRH-RGGSVTVYCGSLISPCWVIGATHCFIDYPKKEDYIVVLGRS 214
Db 316 FADIASHPQAAIFAKHRRSPGERFLCGGILLSSCWILSAAHCFQFRPPHLLTILGRT 375

QY 215 RLNSNTGEMKFEVENLILHKDYSADTLAHNDIALLKIRSKRCAQPSRTIOTICLPS 274
Db 376 YRVVPGEEQKFEVEKIVHKEFDDT--YNDIALLLQKSSRCAQESSVVRTVCLPP 433
QY 275 MYNDPQGTSCBITGFKENSTDYLPBOLKMTVVKLISHRECOQPHYGVSEVTTMCLCA 334
Db 434 ADLQLPDWTCELSGYKHEALSPFYSERLKEARVLYPSSRCTSQHLLNRTVTDNMLCA 493
QY 335 AD-----POWKT-DSCOGDSGGPLVCSLOGRMILTGTIVSNGRCALKDKFQVTVRVSHFL 388
Db 494 GDTSGGFPQANLHDACQDSGGPLVCLNDGRMTLVGLISWGLGCGQKDVGVTVKVTNYL 553
QY 389 PWIRSHTK 396
Db 554 DWIRDNR 561
RESULT 35
US-10-411-037-26
; Sequence 26, Application US/10411037
; Publication No. US2004004346A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bove, Caryn
; TITLE OF INVENTION: ALPHA GALACTOSIDASE A: REMODELING AND GLYCOCONJUGATION OF ALPHA
; FILE OF INVENTION: GALACTOSIDASE A
; FILE REFERENCE: 040853-01-5082
; CURRENT APPLICATION NUMBER: US/10/411,037
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-037-26
Query Match 38.0%; Score 858.5; DB 12; Length 562;
Best Local Similarity 37.5%; Pred. No. 4.9e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGHCIDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSPRCFNGGTCQALYFSDF-VCQCPGFGAGKCCBIDTRATCYEDQGISY 135
QY 59 RKGASTDTMGRCLPWNSTVLTQTYHAHRSDALQGLGKHNCRNPNRPRKCYQVQ 118
Db 136 RGTWSTAESGAECTNWNSSALAKPKYSGRRPDAIRLGLGNHNYCRNPDSDSKPVCYFKA 195
QY 119 LKPLVQECMVHDCADG----- 134
Db 196 GKYSSFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLRWNSMILIGKVTYQNP 255
QY 135 -----KLKQ-----CG-QKTLRPRFKIIGGE 155

196 GKYSSEFCSTPACSEGNDCVFCNGSAYRGTHSLTESGASCLPWNMSILIGKVYTAQNP 255
135 -----KLKFO-----CG-QKTLRPRFKIIG 155
256 AQALGLGKHNYCRNPDGAKPWCHVLKNRLTWBYCDVPSCSTCGLRQYSQPFRIKGG 315
156 FTTIENQWFAIYRRH-RGGSVTVCCGSLISPCWVISATHCFIDYPKKEDYIVVLGRS 214
316 FADIASHPWQAIAFAKRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGT 375
215 RLNSNTGEMKFEVENLILHKDYSADTLAHNDIALLKIRSKRCAQPSRTIOTICLPS 274
376 YRVVPGEEQKFEVEKYIVHKEFDDDT--YDNDIALLOKSDSRCAQSSVVRTVCLPP 433
275 MYNDPQGTSCITGFKENSTDYLYPEQLKMTVVKLIHSHRECQPHYGVSEVTTKMLCA 334
434 ADLQLPDWTCELSGYKHEALSPFYSERLKEAHRVLPSSRCTSQHLLNRTVTDNMLCA 493
335 AD-----POWKT-DSCOGDSGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFL 388
494 GDRSGGPOANLHDACOGDSGGPLVCLNDGRMTLVGLIISWGLCGQKQDVGVTYKVTNYL 553
389 PWIRSHTK 396
554 DWIRDNR 561
RESULT 39
US-10-410-962-26
; Sequence 26, Application US/10410962
; Publication No. US20040077836A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DePrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bowe, Caryn
; TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR: REMODELING AND
; FILE REFERENCE: 040853-01-5054
; CURRENT APPLICATION NUMBER: US/10/410,962
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 26
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-410-962-26
Query Match 38.0%; Score 858.5; DB 16; Length 562;
Best Local Similarity 37.5%; Pred. No. 4.9e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
3 ELHQP-SNCD--CLNGGTCVSNKYPSNIHWCNCPKFGGHCCEIDKSKTCYEGNGHFY 58
77 QCHSVPVKSCSEPRCFNGTCCQALYFSDP-VCOCPGPFAGKCEIDTRATCYEDQGISY 135
59 RGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPMCVYQVG 118
136 RGTWSTAESGAECTWNSSALAKQPSYSGRRPDALRLGLGNHNYCRNPNDRSKPWCYVFXA 195
119 LKPLVQECWHDCA DG----- 134
196 KYISSEFCSTPACSEGNDCVFCNGSAYRGTHSLTESGASCLPWNMSILIGKVYTAQNP 255
135 -----KLKFO-----CG-QKTLRPRFKIIG 155
256 AQALGLGKHNYCRNPDGAKPWCHVLKNRLTWBYCDVPSCSTCGLRQYSQPFRIKGG 315
156 FTTIENQWFAIYRRH-RGGSVTVCCGSLISPCWVISATHCFIDYPKKEDYIVVLGRS 214
316 FADIASHPWQAIAFAKRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGT 375

215 RLNSNTGEMKFEVENLILHKDYSADTLAHNDIALLKIRSKRCAQPSRTIOTICLPS 274
376 YRVVPGEEQKFEVEKYIVHKEFDDDT--YDNDIALLOKSDSRCAQSSVVRTVCLPP 433
275 MYNDPQGTSCITGFKENSTDYLYPEQLKMTVVKLIHSHRECQPHYGVSEVTTKMLCA 334
434 ADLQLPDWTCELSGYKHEALSPFYSERLKEAHRVLPSSRCTSQHLLNRTVTDNMLCA 493
335 AD-----POWKT-DSCOGDSGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFL 388
494 GDRSGGPOANLHDACOGDSGGPLVCLNDGRMTLVGLIISWGLCGQKQDVGVTYKVTNYL 553
389 PWIRSHTK 396
554 DWIRDNR 561
RESULT 39
US-10-410-962-26
; Sequence 26, Application US/10410962
; Publication No. US20040077836A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DePrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bowe, Caryn
; TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR: REMODELING AND
; FILE REFERENCE: 040853-01-5054
; CURRENT APPLICATION NUMBER: US/10/410,962
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 26
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-410-962-26
Query Match 38.0%; Score 858.5; DB 16; Length 562;
Best Local Similarity 37.5%; Pred. No. 4.9e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
3 ELHQP-SNCD--CLNGGTCVSNKYPSNIHWCNCPKFGGHCCEIDKSKTCYEGNGHFY 58
77 QCHSVPVKSCSEPRCFNGTCCQALYFSDP-VCOCPGPFAGKCEIDTRATCYEDQGISY 135
59 RGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPMCVYQVG 118
136 RGTWSTAESGAECTWNSSALAKQPSYSGRRPDALRLGLGNHNYCRNPNDRSKPWCYVFXA 195
119 LKPLVQECWHDCA DG----- 134
196 KYISSEFCSTPACSEGNDCVFCNGSAYRGTHSLTESGASCLPWNMSILIGKVYTAQNP 255
135 -----KLKFO-----CG-QKTLRPRFKIIG 155

256 AQAALGLGKHNYCRNPDGDAKWPCHLVKNRRLTWECYDVPSCSTCGLRQYSQPOFRIKGGJ 315
156 FTTIENQWFAAIVRRH-RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGES 214
316 FADIASHPWQAIAFAKRRSPGERFICGGLISSCWILSAHCFQERFPHHLTVILGRT 375
215 RLNSNTGEMKFEVENILHKOVSADTLAHNDIALKIRSKRGCAQPSRTIQTICLPS 274
376 YRVVPGEEQKFEVEKIYVHKFPDDT--YNDNIALQLKSDSRCAQSSVVRTVCLPP 433
275 MYNDPQGTSCETIGFGENSTDYLYPEQLKMTVVKLISHRECOQHYGSEVTTMCA 334
434 ADLQLPWTCELSGYKHEALSPFYSERLKEAHRVLYPSSRCTSQHLLNRTVTDNMLCA 493
335 AD-----PQWKT-DSCQDGGPLVCSLQGRMTLTGIVSMGRGCAKDKPGVYTRVSHFL 388
494 GDRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDPGVYTKVNYL 553
389 PWIRSHTK 396
554 DWIRDNR 561

ULT 40
10-411-049-26
Sequence 26, Application US/10411049
Publication No. US20040082026A1
GENERAL INFORMATION:
APPLICANT: Neose Technologies, Inc.
APPLICANT: Defrees, Shawn
APPLICANT: Zopf, David
APPLICANT: Bayer, Robert
APPLICANT: Hakes, David
APPLICANT: Chen, Xi
APPLICANT: Howe, Caryn
TITLE OF INVENTION: INTERFERON ALPHA: REMODELING AND GLYCOCONJUGATION OF INTERFERON
TITLE OF INVENTION: ALPHA
FILE REFERENCE: 040853-01-5055
CURRENT APPLICATION NUMBER: US/10/411,049
CURRENT FILING DATE: 2003-04-09
PRIOR APPLICATION NUMBER: US 60/328,523
PRIOR FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 60/344,692
PRIOR FILING DATE: 2001-10-19
PRIOR APPLICATION NUMBER: US 60/387,292
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: US 60/391,777
PRIOR FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: US 60/396,594
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: US 60/404,249
PRIOR FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: US 60/407,527
PRIOR FILING DATE: 2002-08-28
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn version 3.2
SEQ ID NO 26
LENGTH: 562
TYPE: PRT
ORGANISM: Homo sapiens
10-411-049-26

Query Match 38.08; Score 858.5; DB 16; Length 562;
Best Local Similarity 37.58; Pred. No. 4.9e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
3 ELHQVP--SNCD---CLNGTCVSNKYFSNIHWNCNCPKFGQHCEDKSKTCYEGNGHFY 58
77 QCHSVPVKSCSEPCFNGTCCQALYFSDF-VQCEPGFAGKCCEDTATCYEDQGISY 135
59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVG 118

Db 136 RGTWSTAEGSAECTNWNSSALLAQKPYSGRRPDIAIRLGLGNHNYCRNPDNRDPKWPYVKA 195
QY 119 LKPLVQECMVHDCADG----- 134
Db 196 KYSSSEFCSTPACSEGNSSDCYFGNGSAVRGTHSILTESGASCLPWNMSILLIKVYTAQNP 255
QY 135 -----KLKFO-----CG-QKTLRPRFKIIGGE 155
Db 256 AQAALGLGKHNYCRNPDGDAKWPCHLVKNRRLTWECYDVPSCSTCGLRQYSQPOFRIKGGJ 315
QY 156 FTTIENQWFAAIVRRH-RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGES 214
Db 316 FADIASHPWQAIAFAKRRSPGERFICGGLISSCWILSAHCFQERFPHHLTVILGRT 375
QY 215 RLNSNTGEMKFEVENILHKOVSADTLAHNDIALKIRSKRGCAQPSRTIQTICLPS 274
Db 376 YRVVPGEEQKFEVEKIYVHKFPDDT--YNDNIALQLKSDSRCAQSSVVRTVCLPP 433
QY 275 MYNDPQGTSCETIGFGENSTDYLYPEQLKMTVVKLISHRECOQHYGSEVTTMCA 334
Db 434 ADLQLPWTCELSGYKHEALSPFYSERLKEAHRVLYPSSRCTSQHLLNRTVTDNMLCA 493
QY 335 AD-----PQWKT-DSCQDGGPLVCSLQGRMTLTGIVSMGRGCAKDKPGVYTRVSHFL 388
Db 494 GDRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDPGVYTKVNYL 553
QY 389 PWIRSHTK 396
Db 554 DWIRDNR 561

RESULT 41
US-09-880-503-4
Sequence 4, Application US/09880503
Patent No. US20020131984A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
TITLE OF INVENTION: TISSUE CONTRACTABILITY
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 135
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-4

Query Match 35.14; Score 793; DB 9; Length 135;
Best Local Similarity 100.08; Pred. No. 5.1e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPNSCDCLNGTCTVSNKYFSNIHWNCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
Db 1 SNELHQVPNSCDCLNGTCTVSNKYFSNIHWNCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135

RESULT 42
US-09-984-186-12
Sequence 12, Application US/09984186

Patent No. US20020151011A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-984-186-12
Query Match 35.1%; Score 793; DB 9; Length 138;
Best Local Similarity 100.0%; Pred. No. 5.3e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNTHWCNCPKKGQCHCEIDKSKTCYEGNGHFYRG 60
DE 4 SNELHQPNSCCLNGGTCVSNKYFSNTHWCNCPKKGQCHCEIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 120
DL 64 KASTDTMGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DL 124 PLVQECMVHDCADGK 138
RE LT 43
US 0-237-667-12
; quence 12, Application US/10237667
; ublication No. US20030022308A1
; GENERAL INFORMATION:

GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,667
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-667-12
Query Match 35.1%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 5.3e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNTHWCNCPKKGQCHCEIDKSKTCYEGNGHFYRG 60
DB 4 SNELHQPNSCCLNGGTCVSNKYFSNTHWCNCPKKGQCHCEIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 120
DB 64 KASTDTMGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138
RESULT 44
US-10-237-708-12
; Sequence 12, Application US/10237708
; Publication No. US20030036170A1
; GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,708
FILING DATE: 10-Sep-2002
CLASSIFICATION: <UNKNOWN>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
10-237-708-12

Query Match 35.1%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 5.3e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
D 4 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 63

Q 61 KASDTMTGRPCLPWNSATVLQOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
D 64 KASDTMTGRPCLPWNSATVLQOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 123

C 121 PLVQECMVHDCADGK 135
I 124 PLVQECMVHDCADGK 138

RESULT 45
US-10-237-866-12
Sequence 12, Application US/10237866
Publication No. US20030036171A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard

Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <UNKNOWN>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match 35.1%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 5.3e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
D 4 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 63

Q 61 KASDTMTGRPCLPWNSATVLQOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
D 64 KASDTMTGRPCLPWNSATVLQOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 123

Q 121 PLVQECMVHDCADGK 135
D 124 PLVQECMVHDCADGK 138

RESULT 46
US-10-237-871-12
Sequence 12, Application US/10237871
Publication No. US20030036172A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain

Guillon, Jean-Dominique
Jung, Gerard
Yeh, Patricia
Yeh, Patricia
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,871
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-0-237-871-12
Query Match 35.1%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 5.3e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRCPLPWSATVLTQTYHAHRSALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 64 KASTDTMGRCPLPWSATVLTQTYHAHRSALQGLGKHNYCRPNDRRPPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138
RESULT 48
US-10-702-536-12
Sequence 12, Application US/10702536
Publication NO. US20040086976A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guillon, Jean-Dominique
Jung, Gerard

Guillon, Jean-Dominique
Jung, Gerard
Yeh, Patricia
Yeh, Patricia
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,871
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-0-237-871-12
Query Match 35.1%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 5.3e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRCPLPWSATVLTQTYHAHRSALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 64 KASTDTMGRCPLPWSATVLTQTYHAHRSALQGLGKHNYCRPNDRRPPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138
RESULT 47
US-0-237-624-12
Sequence 12, Application US/10237624
Publication NO. US20030082747A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guillon, Jean-Dominique
Jung, Gerard

Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/702,536
FILING DATE: 07-Nov-2003
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US

TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-702-536-12

Query Match 35.1%; Score 793; DB 16; Length 138;
Best Local Similarity 100.0%; Pred. No. 5.3e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHELHQVPSNCCLNGGTCVSNKYFSNIHWNCNPKFGQHCIEDKSKTCYEGNGHYRG 60
DB 4 SHELHQVPSNCCLNGGTCVSNKYFSNIHWNCNPKFGQHCIEDKSKTCYEGNGHYRG 63
QY 61 KASTDTMGPRCLPWNASATVLQOTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
DB 64 KASTDTMGPRCLPWNASATVLQOTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 50
US-09-880-503-8
; Sequence 8, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503

```

; CURRENT FILING DATE: 2001-06-13
; RIOR APPLICATION NUMBER: US 60/212,847
; RIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; 2 ID NO 8
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
; 9-880-503-8

```

```

37y Match 35.13; Score 793; DB 9; Length 143;
st Local Similarity 100.0%; Pred. No. 5.5e-62;
tches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DE 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWNSATVLOOTYHAHRS DALQLGLGKHN YCRNPDNRREPWCYQVGLK 120
DE 61 KASTDTMGRPCLPWNSATVLOOTYHAHRS DALQLGLGKHN YCRNPDNRREPWCYQVGLK 120

QY 121 PLVQECMVHDCADGK 135
DE 121 PLVQECMVHDCADGK 135

```

```

Se ch completed: May 25, 2004, 15:03:47
Jc time : 76.4654 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

(protein - protein search, using sw model

on: May 25, 2004, 14:48:05 ; Search time 27.9027 Seconds
(without alignments)
745.636 Million cell updates/sec

le: US-09-880-503-6

fect score: 2257

uence: 1 SNEHQVPSNCDLNGTGV.....VSHFLPWIRSHTKBENGALL 403

ring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 389414 seqs, 51625971 residues

al number of hits satisfying chosen parameters: 389414

imum DB seq length: 0

imum DB seq length: 2000000000

t-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

abase : Issued Patents AA:*

1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*

2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*

3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*

4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*

5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pcp.*

6: /cgn2_6/ptodata/2/iaa/backfiles.pcp.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

ult No.	Score	Query Match	Length	DB ID	Description
1	2243.5	99.4	430	6	Patent No. 5219569-2
2	2243	99.4	411	1	Sequence 1, Appl
3	2243	99.4	411	1	Sequence 18, Appl
4	2243	99.4	411	1	Sequence 18, Appl
5	2243	99.4	431	4	Sequence 1, Appl
6	2243	99.4	431	6	Patent No. 5188829
7	2240	99.2	411	4	Sequence 2, Appl
8	2240	99.2	430	1	Sequence 3, Appl
9	2233	98.9	411	3	Sequence 1, Appl
10	2219	98.3	411	2	Sequence 48, Appl
11	1965	87.1	432	2	Sequence 47, Appl
12	1964	87.0	365	1	Sequence 83, Appl
13	1964	87.0	365	1	Sequence 83, Appl
14	1964	87.0	393	2	Sequence 24, Appl
15	1964	87.0	393	3	Sequence 24, Appl
16	1964	87.0	393	3	Sequence 25, Appl
17	1466.5	85.0	306	2	Sequence 45, Appl
18	1466.5	85.0	331	2	Sequence 46, Appl
19	1382	61.2	253	3	Sequence 73, Appl
20	1374	60.9	254	2	Sequence 49, Appl
21	983	43.6	200	4	Sequence 73, Appl
22	867.5	38.4	477	2	Sequence 51, Appl
23	865.5	38.3	527	2	Sequence 39, Appl
24	859.5	38.1	527	2	Patent No. 5520313
25	858.5	38.0	527	1	Sequence 16, Appl
26	858.5	38.0	527	5	Sequence 2, Appl
27	858.5	38.0	527	6	Patent No. 5185259

28	858.5	38.0	562	2	US-08-811-949-43	Sequence 43, Appl
29	858.5	38.0	562	2	US-08-560-098A-30	Sequence 50, Appl
30	858.5	38.0	562	2	US-08-883-795A-58	Sequence 38, Appl
31	858.5	38.0	562	4	US-09-703-695A-4	Sequence 4, Appl
32	858.5	38.0	562	6	5185259-3	Patent No. 5185259
33	858.5	38.0	562	6	5200340-2	Patent No. 5200340
34	858.5	38.0	562	6	5344773-2	Patent No. 5344773
35	851.5	37.7	562	6	5244676-5	Patent No. 5244676
36	825	35.6	157	3	US-08-142-590B-25	Sequence 25, Appl
37	808	35.8	355	2	US-08-811-949-59	Sequence 59, Appl
38	803	35.6	437	2	US-08-811-949-51	Sequence 51, Appl
39	800	35.4	437	2	US-08-811-949-57	Sequence 57, Appl
40	799	35.4	208	4	US-09-101-272G-98	Sequence 98, Appl
41	796	35.3	472	2	US-08-811-949-63	Sequence 63, Appl
42	793	35.1	138	2	US-08-797-689-12	Sequence 12, Appl
43	793	35.1	138	4	US-09-984-186-12	Sequence 12, Appl
44	793	35.1	437	2	US-08-811-949-55	Sequence 55, Appl
45	792	35.1	437	2	US-08-811-949-49	Sequence 49, Appl
46	791	35.0	355	2	US-08-811-949-47	Sequence 47, Appl
47	790	35.0	356	1	US-08-427-640-8	Sequence 8, Appl
48	788	34.9	194	4	US-09-101-272G-80	Sequence 80, Appl
49	788	34.9	201	4	US-09-101-272G-96	Sequence 96, Appl
50	786	34.8	378	4	US-09-553-498-10	Sequence 10, Appl
51	786	34.8	378	4	US-09-618-869-10	Sequence 10, Appl
52	784	34.7	355	1	US-08-137-116-1	Sequence 1, Appl
53	784	34.7	355	1	US-08-217-618-1	Sequence 1, Appl
54	784	34.7	355	1	US-08-427-640-2	Sequence 2, Appl
55	784	34.7	355	1	US-08-217-616-1	Sequence 1, Appl
56	784	34.7	355	1	US-08-217-616-1	Sequence 1, Appl
57	784	34.7	355	2	US-08-811-949-45	Sequence 45, Appl
58	784	34.7	355	3	US-08-794-528-1	Sequence 1, Appl
59	784	34.7	355	6	5232356-1	Patent No. 5232356
60	781	34.6	355	2	US-08-811-949-53	Sequence 53, Appl
61	780	34.6	389	2	US-08-811-949-67	Sequence 67, Appl
62	779	34.5	355	1	US-08-427-640-6	Sequence 6, Appl
63	779	34.5	389	2	US-08-811-949-65	Sequence 65, Appl
64	769.5	34.1	356	1	US-08-427-640-4	Sequence 4, Appl
65	766.5	34.0	354	2	US-08-811-949-61	Sequence 61, Appl
66	765.5	33.9	347	2	US-08-811-949-1	Sequence 1, Appl
67	726.5	32.2	655	1	US-08-148-910-12	Sequence 12, Appl
68	726.5	32.2	655	1	US-08-448-937A-12	Sequence 12, Appl
69	645	28.6	546	6	5200340-6	Patent No. 5200340
70	633.5	28.1	326	4	US-09-411-977-3	Sequence 3, Appl
71	593	25.8	383	3	US-08-558-269-6	Sequence 6, Appl
72	583	25.8	383	3	US-09-410-882-6	Sequence 6, Appl
73	565.5	25.1	253	2	US-09-027-337-8	Sequence 8, Appl
74	565.5	25.1	253	4	US-09-644-600-8	Sequence 8, Appl
75	565.5	25.1	253	4	US-09-654-600A-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
5219569-2
; Patent No. 5219569
; APPLICANT: ELABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO:2:
; LENGTH: 430
5219569-2

Query Match 99.4%; Score 2243.5; DB 6; Length 430;
Best Local Similarity 98.3%; Pred. No. 4.3e-188;
Matches 403; Conservative 0; Mismatches 0; Indels 7; Gaps 1;

1 SNELHVPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 60
 21 SNELHVPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 80
 61 KASDTTMRGPRCLPNSATVLOQTYHAHRSALQGLGKHNYCRPNDRRRPWCYVQVGLK 120
 81 KASDTTMRGPRCLPNSATVLOQTYHAHRSALQGLGKHNYCRPNDRRRPWCYVQVGLK 140
 121 PLVQECMVHDCADGK-----LAFQCGOKTLRPRFKIIGGEFTTIENQWFAAIYRRH 173
 141 PLVQECMVHDCADGKPPPELAFQCGOKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200
 174 GGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFEVENLIL 233
 201 GGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFEVENLIL 260
 234 HKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGK 293
 261 HKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGK 320
 294 NSTDYLYPEQLKMTVVKLIISHRECCQPHYGVSEVTTKMLCAADPWKTDSCQDSDGGLV 353
 321 NSTDYLYPEQLKMTVVKLIISHRECCQPHYGVSEVTTKMLCAADPWKTDSCQDSDGGLV 380
 354 CSLQGRMTLTGIVSWGRGKALDKRPGVYTVRVSHFLPWIRSHTKENGIAL 403
 381 CSLQGRMTLTGIVSWGRGKALDKRPGVYTVRVSHFLPWIRSHTKENGIAL 430

RE LT 2
 US 8-087-163-1
 ; Sequence 1, Application US/08087163
 ; Patent No. 5472692
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jian-Ning
 ; TITLE OF INVENTION: PRO-UKINASE MUTANTS
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 502 or 55SX
 ; OPERATING SYSTEM: MS-DOS (Version 5.0)
 ; SOFTWARE: WordPerfect (Version 5.1)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/087,163
 ; FILING DATE: 07/02/93
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fasse, J. Peter
 ; REGISTRATION NUMBER: 32,983
 ; REFERENCE/DOCKET NUMBER: 04353/003001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 542-5070
 ; TELEFAX: (617) 542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 411
 ; TYPE: amino acid
 ; STRANDEDNESS: N/A
 ; TOPOLOGY: N/A
 ; 8-087-163-1

Query Match 99.4%; Score 2243; DB 1; Length 411;
 Best Local Similarity 98.1%; Pred. No. 4.4e-188;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNELHVPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 60
 DB 1 SNELHVPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 60
 QY 61 KASDTTMRGPRCLPNSATVLOQTYHAHRSALQGLGKHNYCRPNDRRRPWCYVQVGLK 120
 DB 61 KASDTTMRGPRCLPNSATVLOQTYHAHRSALQGLGKHNYCRPNDRRRPWCYVQVGLK 120
 QY 121 PLVQECMVHDCADGK-----LAFQCGOKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
 DB 121 PLVQECMVHDCADGKPPPELAFQCGOKTLRPRFKIIGGEFTTIENQWFAAIYRRH 180
 QY 173 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFEVENLIL 232
 DB 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFEVENLIL 240
 QY 233 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGK 292
 DB 241 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGK 300
 QY 293 ENSTDYLYPEQLKMTVVKLIISHRECCQPHYGVSEVTTKMLCAADPWKTDSCQDSDGGLV 352
 DB 301 ENSTDYLYPEQLKMTVVKLIISHRECCQPHYGVSEVTTKMLCAADPWKTDSCQDSDGGLV 360
 QY 353 VCSLQGRMTLTGIVSWGRGKALDKRPGVYTVRVSHFLPWIRSHTKENGIAL 403
 DB 361 VCSLQGRMTLTGIVSWGRGKALDKRPGVYTVRVSHFLPWIRSHTKENGIAL 411

RESULT 3
 US-08-286-748B-18
 ; Sequence 18, Application US/08286748B
 ; Patent No. 5759542
 ; GENERAL INFORMATION:
 ; APPLICANT: Victor Gurewich
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
 ; TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
 ; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 502 or 55SX
 ; OPERATING SYSTEM: MS-DOS (Version 5.0)
 ; SOFTWARE: WordPerfect (Version 5.1)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/286,748B
 ; FILING DATE: August 5, 1994
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: J. Peter Fasse
 ; REGISTRATION NUMBER: 32,983
 ; REFERENCE/DOCKET NUMBER: 04547/013001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 542-5070
 ; TELEFAX: (617) 542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:


```

OTHER INFORMATION:
NAME/KEY: misc.feature
LOCATION: (20)...()
OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)
9-101-272G-1

ary Match 99.4%; Score 2243; DB 4; Length 431;
st Local Similarity 98.1%; Pred. No. 4.7e-188;
tches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPVSNCDCNLGGTGVSNKYPFNSNHCNCPKFGGQHCEDKSKTCYEGNGHFYRG 60
Df 21 SNELHQPVSNCDCNLGGTGVSNKYPFNSNHCNCPKFGGQHCEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
Df 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQGQKTLRPRFKIIGGEBFTTIENQPFWFAAIYRRH 172
Df 141 PLVQECMVHDCADGKPKSSPPEELKFQCGQKTLRPRFKIIGGEBFTTIENQPFWFAAIYRRH 200
QY 173 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 232
Df 201 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALLKIRSKGRCALDKPGVYTVSVSHFLPWIRSHTKBEENGLAL 292
Df 261 LHKDYSADTLAHHNDIALLKIRSKGRCALDKPGVYTVSVSHFLPWIRSHTKBEENGLAL 320
QY 293 ENSTDYLYPEOLKMTVVKVLIHSHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDSGGPL 352
Df 321 ENSTDYLYPEOLKMTVVKVLIHSHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDSGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRGKALDKPGVYTVSVSHFLPWIRSHTKBEENGLAL 403
Df 381 VCSLQGRMTLTGIVSWGRGKALDKPGVYTVSVSHFLPWIRSHTKBEENGLAL 431

RE UT 6
51 929-1
; Applicant: KOBAYASHI, YO-ICHI, OMORI, MUNEKI, YAMADA, CHIKAKO
; TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1988
; ID NO.1:
; LENGTH: 431.
51 829-1

ery Match 99.4%; Score 2243; DB 6; Length 431;
st Local Similarity 98.1%; Pred. No. 4.7e-188;
tches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPVSNCDCNLGGTGVSNKYPFNSNHCNCPKFGGQHCEDKSKTCYEGNGHFYRG 60
Df 21 SNELHQPVSNCDCNLGGTGVSNKYPFNSNHCNCPKFGGQHCEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
Df 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQGQKTLRPRFKIIGGEBFTTIENQPFWFAAIYRRH 172
Df 141 PLVQECMVHDCADGKPKSSPPEELKFQCGQKTLRPRFKIIGGEBFTTIENQPFWFAAIYRRH 200
QY 173 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 232
Df 201 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALLKIRSKGRCALDKPGVYTVSVSHFLPWIRSHTKBEENGLAL 292

```

```

Db 261 LHKDYSADTLAHHNDIALLKIRSKGRCALDKPGVYTVSVSHFLPWIRSHTKBEENGLAL 431
QY 293 ENSTDYLYPEOLKMTVVKVLIHSHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDSGGPL 352
Df 321 ENSTDYLYPEOLKMTVVKVLIHSHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDSGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRGKALDKPGVYTVSVSHFLPWIRSHTKBEENGLAL 403
Df 381 VCSLQGRMTLTGIVSWGRGKALDKPGVYTVSVSHFLPWIRSHTKBEENGLAL 431

RESULT 7
US-09-403-736-2
; Sequence 2, Application US/09403736
; Patent No. 6638502
; GENERAL INFORMATION:
; APPLICANT: Aventis S.A.
; APPLICANT: LI, Hong
; APPLICANT: LU, He
; APPLICANT: GRISCELLI, Frank
; APPLICANT: OPOLON, Paule
; APPLICANT: SORIA, Claudine
; APPLICANT: RAGOT, Thierry
; APPLICANT: LEGRAND, Yves
; APPLICANT: SORIA, Jeanette
; APPLICANT: MABILLAT, Christelle
; APPLICANT: FERRICAUDET, Michel
; APPLICANT: YEH, Patrice
; TITLE OF INVENTION: Adenovirus-Mediated Intratumoral Delivery Of An Angiogenesis Anta
; FILE REFERENCE: A2778A-US
; CURRENT APPLICATION NUMBER: US/09/403,736
; CURRENT FILING DATE: 1999-10-26
; PRIOR APPLICATION NUMBER: PCT/EP98/02491
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/044,980
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: humanurokinase
US-09-403-736-2

Query Match 99.2%; Score 2240; DB 4; Length 411;
Best Local Similarity 97.8%; Pred. No. 8.1e-185;
Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPVSNCDCNLGGTGVSNKYPFNSNHCNCPKFGGQHCEDKSKTCYEGNGHFYRG 60
Df 1 SNELHQPVSNCDCNLGGTGVSNKYPFNSNHCNCPKFGGQHCEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
Df 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK-----LKFCQGQKTLRPRFKIIGGEBFTTIENQPFWFAAIYRRH 172
Df 121 PLVQECMVHDCADGKPKSSPPEELKFQCGQKTLRPRFKIIGGEBFTTIENQPFWFAAIYRRH 180
QY 173 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 232
Df 181 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 240
QY 233 LHKDYSADTLAHHNDIALLKIRSKGRCALDKPGVYTVSVSHFLPWIRSHTKBEENGLAL 292
Df 241 LHKDYSADTLAHHNDIALLKIRSKGRCALDKPGVYTVSVSHFLPWIRSHTKBEENGLAL 300
QY 293 ENSTDYLYPEOLKMTVVKVLIHSHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDSGGPL 352
Df 301 ENSTDYLYPEOLKMTVVKVLIHSHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDSGGPL 360

```


US 8-560-098A-48
 ; quence 48; Application US/08560098A
 ; Patent No. 5976841
 ; GENERAL INFORMATION:
 ; APPLICANT: WNEUDT, Stephan
 ; APPLICANT: HEINZEL-WIELAND, Regina
 ; APPLICANT: STEFFENS, Gerd Josef
 ; TITLE OF INVENTION: Proteins having Fibrinolytic and
 ; TITLE OF INVENTION: Coagulation-inhibiting Properties
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 ; STREET: 1200 G Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/560,098A
 ; FILING DATE: 17-NOV-1995
 ; PRIOR APPLICATION NUMBER: P 44 40 892.7
 ; FILING DATE: 17-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: EVANS, Joseph D.
 ; REGISTRATION NUMBER: 26,269
 ; REFERENCE/DOCKET NUMBER: 148/42448
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-8800
 ; TELEFAX: (202) 628-8844
 ; INFORMATION FOR SEQ ID NO: 48:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 411 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; 8-560-098A-48

ery Match 98.3%; Score 2219; DB 2; Length 411;
 st Local Similarity 97.3%; Pred. No. 5.6e-186;
 tches 400; Conservative 0; Mismatches 3; Indels 8; Gaps 1;

QY 1 SNELHVPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
 Df 1 SNELHVPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
 QY 61 KASDTMGRPCLPNSATVLOQTHAHRSDALQLGLGKHNCRPNRRPWCYVQVGLK 120
 Df 61 KASDTMGRPCLPNSATVLOQTHAHRSDALQLGLGKHNCRPNRRPWCYVQVGLK 120
 QY 121 PLVCEMWHDCADGK-----LKFCGQKTLRPFKIIIGGFTTIENQFWFAALYRRH 172
 Df 121 PLVCEMWHDCADGKSPPEELKFCGQKTLRPFKIIIGGFTTIENQFWFAALYRRH 180
 QY 173 RGGSVTVVCGSLSPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 232
 Df 181 RGGSVTVVCGSLSPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 240
 QY 233 LHKDYSADTLAHDNDIALLKIRSEGRCAQPSRTIOTICLPMSYNDPQFGTSCITGFGK 292
 Df 241 LHKDYSADTLAHDNDIALLKIRSEGRCAQPSRTIOTICLPMSYNDPQFGTSCITGFGK 300
 QY 293 ENSDYLYPEOLKMTVVKLISHRECQPHYGVSEVTVMKLCADDPQWKTSCOGDSGGPL 352
 Df 301 ENSDYLYPEOLKMTVVKLISHRECQPHYGVSEVTVMKLCADDPQWKTSCOGDSGGPL 360
 QY 353 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGAL 403

Db 361 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGAL 411
 RESULT 11
 US-08-560-098A-47
 ; Sequence 47; Application US/08560098A
 ; Patent No. 5976841
 ; GENERAL INFORMATION:
 ; APPLICANT: WNEUDT, Stephan
 ; APPLICANT: HEINZEL-WIELAND, Regina
 ; APPLICANT: STEFFENS, Gerd Josef
 ; TITLE OF INVENTION: Proteins having Fibrinolytic and
 ; TITLE OF INVENTION: Coagulation-inhibiting Properties
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 ; STREET: 1200 G Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/560,098A
 ; FILING DATE: 17-NOV-1995
 ; PRIOR APPLICATION NUMBER: P 44 40 892.7
 ; FILING DATE: 17-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: EVANS, Joseph D.
 ; REGISTRATION NUMBER: 26,269
 ; REFERENCE/DOCKET NUMBER: 148/42448
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-8800
 ; TELEFAX: (202) 628-8844
 ; INFORMATION FOR SEQ ID NO: 47:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 432 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-560-098A-47

Query Match 87.1%; Score 1965; DB 2; Length 432;
 Best Local Similarity 90.5%; Pred. No. 9.5e-164;
 Matches 361; Conservative 8; Mismatches 16; Indels 14; Gaps 2;

QY 13 CLNGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRGKASTDTMGRPCL 72
 Df 40 CVTGEPTKPSHNGDFFEEIPEEY-----LQISKTCEYEGNGHYRGKASTDTMGRPCL 93
 QY 73 PWSATVLOQTHAHRSDALQLGLGKHNCRPNRRPWCYVQVGLKPLVCEMWHDC 132
 Df 94 PWSATVLOQTHAHRSDALQLGLGKHNCRPNRRPWCYVQVGLKPLVCEMWHDC 153
 QY 133 DCK-----LKFCGQKTLRPFKIIIGGFTTIENQFWFAALYRRHGGSVTVVCGGS 184
 Df 154 DCKFPSPPEELKFCGQKTLRPFKIIIGGFTTIENQFWFAALYRRHGGSVTVVCGGS 213
 QY 185 LISPQWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLIHKDYADTLAH 244
 Df 214 LISPQWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLIHKDYADTLAH 273
 QY 245 HNDIALLKIRSEGRCAQPSRTIOTICLPMSYNDPQFGTSCITGFGKENSTDYLYPEOL 304
 Df 274 HNDIALLKIRSEGRCAQPSRTIOTICLPMSYNDPQFGTSCITGFGKENSTDYLYPEOL 333

C 305 KMTVVKLISHRECOQPHYYGSEVTTKMLCAADPOWKTDSQCGDSGGPLVCSLQGRMTLTG 364
D 334 KMTVVKLISHRECOQPHYYGSEVTTKMLCAADPOWKTDSQCGDSGGPLVCSLQGRMTLTG 393
C 365 IVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 403
D 394 IVSWGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 432
F JUL 12
US-08-093-741-83
Sequence 83, Application US/08093741
Patent No. 5681721
GENERAL INFORMATION:
APPLICANT: STEPPENS, GERT J.
APPLICANT: WENDT, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,741
FILING DATE: 20-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-093-741-83
Query Match 87.0%; Score 1964; DB 1; Length 365;
Best Local Similarity 97.8%; Pred. No. 9.4e-164;
Matches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
C 47 SKTCYEGNGHFYRGKASTDTMGPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPD 106
D 1 SKTCYEGNGHFYRGKASTDTMGPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPD 60
C 107 NRRRPWCYVQVGLKPLVQECMWHDCADGK-----LKFCQCKTLRPFKLIIGSEFTT 158
D 61 NRRRPWCYVQVGLKPLVQECMWHDCADGKPKSPPEELKFQCGQKTLRPFKLIIGSEFTT 120
C 159 IENQPFALYRRHGRGSEVTVYCGGSLISPCWVISAATHCFIDYPKKEDYIVVLGSRRLNS 218
D 121 IENQPFALYRRHGRGSEVTVYCGGSLISPCWVISAATHCFIDYPKKEDYIVVLGSRRLNS 180
C 219 NTQGMKFEVENLILHKDYSATLHAHNDIALLKIRSEGRCAQPSRTIQTICLPSMYND 278

Db 181 NTQGMKFEVENLILHKDYSATLHAHNDIALLKIRSEGRCAQPSRTIQTICLPSMYND 240
QY 279 POFGISCETITGKGNSTDYLYPEQLKMTVVKLISHRECOQPHYYGSEVTTKMLCAADPO 338
Db 241 POFGISCETITGKGNSTDYLYPEQLKMTVVKLISHRECOQPHYYGSEVTTKMLCAADPO 300
QY 339 WKTDSQCGDSGGPLVCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKEE 398
Db 301 WKTDSQCGDSGGPLVCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKEE 360
QY 399 NGLAL 403
Db 361 NGLAL 365
RESULT 13
US-08-720-012-83
Sequence 83, Application US/08720012
Patent No. 5747291
GENERAL INFORMATION:
APPLICANT: STEPPENS, GERT J.
APPLICANT: WENDT, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,012
FILING DATE: 27-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,741
FILING DATE: 20-JUL-1993
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-720-012-83
Query Match 87.0%; Score 1964; DB 1; Length 365;
Best Local Similarity 97.8%; Pred. No. 9.4e-164;
Matches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 47 SKTCYEGNGHFYRGKASTDTMGPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPD 106
Db 1 SKTCYEGNGHFYRGKASTDTMGPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPD 60
QY 107 NRRRPWCYVQVGLKPLVQECMWHDCADGK-----LKFCQCKTLRPFKLIIGSEFTT 158

DE 61 NRRPWCYVQVGLKPLVQECWVHDCADGKPPPEELKFCQGGKTLRPFKIIIGBEFTT 120
QY 159 IENQPFWAIIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGSRILNS 218
DE 121 IENQPFWAIIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGSRILNS 180
QY 219 NTOGEMKFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCAQPSRTIQTICLPSMYND 278
DE 181 NTOGEMKFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCAQPSRTIQTICLPSMYND 240
QY 279 POFGTSCEITGFGKENSIDYLPKQMTVVKLISHRECOQPHYVGSVTTKMLCAADPO 338
DE 241 POFGTSCEITGFGKENSIDYLPKQMTVVKLISHRECOQPHYVGSVTTKMLCAADPO 300
QY 339 WKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGKALDKKPGVYTRVSHFLPWIRSHTKEE 398
DE 301 WKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGKALDKKPGVYTRVSHFLPWIRSHTKEE 360
QY 399 NGLAL 403
DE 361 NGLAL 365

RE LT 14
US 3-560-098A-44
; Sequence 44, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US 3-560-098A-44

ery Match 87.0%; Score 1964; DB 2; Length 393;
st Local Similarity 97.8%; Pred. No. 1e-163;
tches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
US 3-560-098A-44
QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPD 106

Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPD 61
QY 107 NRRPWCYVQVGLKPLVQECWVHDCADGK-----LKFCQGGKTLRPFKIIIGBEFTT 158
Db 62 NRRPWCYVQVGLKPLVQECWVHDCADGKPKSPPEELKFCQGGKTLRPFKIIIGBEFTT 121
QY 159 IENQPFWAIIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGSRILNS 218
Db 122 IENQPFWAIIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGSRILNS 181
QY 219 NTOGEMKFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCAQPSRTIQTICLPSMYND 278
Db 182 NTOGEMKFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCAQPSRTIQTICLPSMYND 241
QY 279 POFGTSCEITGFGKENSIDYLPKQMTVVKLISHRECOQPHYVGSVTTKMLCAADPO 338
Db 242 POFGTSCEITGFGKENSIDYLPKQMTVVKLISHRECOQPHYVGSVTTKMLCAADPO 301
QY 339 WKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGKALDKKPGVYTRVSHFLPWIRSHTKEE 398
Db 302 WKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGKALDKKPGVYTRVSHFLPWIRSHTKEE 361
QY 399 NGLAL 403
Db 362 NGLAL 366

RESULT 15
US-08-967-024C-24
; Sequence 24, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,024C
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 42 665.8
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-967-024C-24

Query Match 87.0%; Score 1964; DB 3; Length 393;
Best Local Similarity 97.8%; Pred. No. 1e-163;
Matches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
US-08-967-024C-25

47 SKTCYEGNGHGYRKGASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPD 106
2 SKTCYEGNGHGYRKGASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPD 61
107 NRRRPWCYVQVGLKPLVQECWHDCAADGK-----LKFCQCGKTLRPRFKIIGGEFTT 158
62 NRRRPWCYVQVGLKPLVQECWHDCAADGK-----LKFCQCGKTLRPRFKIIGGEFTT 121
159 IENQPFPAALYRRHRGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNS 218
122 IENQPFPAALYRRHRGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNS 181
219 NTQGMKFEVENILHKDYSADTLAHNDIALKIRSGRCALDKPGVYTRVSHFLPWIRSHTKEE 278
182 NTQGMKFEVENILHKDYSADTLAHNDIALKIRSGRCALDKPGVYTRVSHFLPWIRSHTKEE 241
279 PQFGTSCIEITGFGKENSVDLYPEQLKMTVVKLIHRECCQPHYVGSVTTKMLCAADPQ 338
242 PQFGTSCIEITGFGKENSVDLYPEQLKMTVVKLIHRECCQPHYVGSVTTKMLCAADPQ 301
339 WKTDSCQSGGGLVCSLOGRMVLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKEE 398
302 WKTDSCQSGGGLVCSLOGRMVLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKEE 361
399 NGLAL 403
362 NGLAL 366

RESULT 16

US-08-967-024C-25
Sequence 25, Application US/08967024C
Patent No. 6133011
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOSHA, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 87.0%; Score 1964; DB 3; Length 393;
Best Local Similarity 97.8%; Pred. No. 1e-163;
Matches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
US-08-967-024C-25

47 SKTCYEGNGHGYRKGASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPD 106
2 SKTCYEGNGHGYRKGASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPD 61
107 NRRRPWCYVQVGLKPLVQECWHDCAADGK-----LKFCQCGKTLRPRFKIIGGEFTT 158
62 NRRRPWCYVQVGLKPLVQECWHDCAADGK-----LKFCQCGKTLRPRFKIIGGEFTT 121
159 IENQPFPAALYRRHRGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNS 218
122 IENQPFPAALYRRHRGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNS 181
219 NTQGMKFEVENILHKDYSADTLAHNDIALKIRSGRCALDKPGVYTRVSHFLPWIRSHTKEE 278
182 NTQGMKFEVENILHKDYSADTLAHNDIALKIRSGRCALDKPGVYTRVSHFLPWIRSHTKEE 241
279 PQFGTSCIEITGFGKENSVDLYPEQLKMTVVKLIHRECCQPHYVGSVTTKMLCAADPQ 338
242 PQFGTSCIEITGFGKENSVDLYPEQLKMTVVKLIHRECCQPHYVGSVTTKMLCAADPQ 301
339 WKTDSCQSGGGLVCSLOGRMVLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKEE 398
302 WKTDSCQSGGGLVCSLOGRMVLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKEE 361
399 NGLAL 403
362 NGLAL 366

RESULT 17

US-08-560-098A-45
Sequence 45, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800

TELEPHONE: (202) 628-8844
 TELEFAX: (202) 628-8844
 INFORMATION FOR SEQ ID NO: 45:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 306 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-560-098A-45

Query Match 65.0%; Score 1466.5; DB 2; Length 306;
 Best Local Similarity 91.1%; Pred. No. 2.3e-120;
 Matches 275; Conservative 6; Mismatches 12; Indels 9; Gaps 3;

103 RNPNNRRP-WCYVQVGLKPLVQECMVHDCADGKLFQCCGQKTLRPRFKLIGGEFTTIE 161
 13 RNPNDKYEPFWEDEKQ-PHMS-----PPELKFQCCGQKTLRPRFKLIGGEFTTIE 64
 162 QPWFPAAYRRHRGGSVTVVCGSLISPCWVISATFCFIDYPKKEDYIVVLRGRSLNSNTQ 221
 65 QPWFPAAYRRHRGGSVTVVCGSLISPCWVISATFCFIDYPKKEDYIVVLRGRSLNSNTQ 124
 222 GEMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQF 281
 125 GEMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQF 184
 282 GTSCEITGFGKENSVDLYPEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPQWKT 341
 185 GTSCEITGFGKENSVDLYPEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPQWKT 244
 342 DSCQDGGGLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGL 401
 245 DSCQDGGGLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGL 304
 402 AL 403
 305 AL 306

US 560-098A-46
 Sequence 46, Application US/08560098A
 Patent No. 5976841
 GENERAL INFORMATION:
 APPLICANT: WENDT, Stephan
 APPLICANT: HEINZEL-WIELAND, Regina
 APPLICANT: STEFFENS, Gerd Josef
 TITLE OF INVENTION: Proteins having Fibrinolytic and
 TITLE OF INVENTION: Coagulation-inhibiting Properties
 NUMBER OF SEQUENCES: 60
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 STREET: 1200 G Street, N.W., Suite 700
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/560,098A
 FILING DATE: 17-NOV-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: P 44 40 892.7
 FILING DATE: 17-NOV-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: EVANS, Joseph D.
 REGISTRATION NUMBER: 26,269
 REFERENCE/DOCKET NUMBER: 148/42448
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800
 TELEFAX: (202) 628-8844
 INFORMATION FOR SEQ ID NO: 46:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 331 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-560-098A-46

Query Match 65.0%; Score 1466.5; DB 2; Length 331;
 Best Local Similarity 91.1%; Pred. No. 2.6e-120;
 Matches 275; Conservative 6; Mismatches 12; Indels 9; Gaps 3;

103 RNPNNRRP-WCYVQVGLKPLVQECMVHDCADGKLFQCCGQKTLRPRFKLIGGEFTTIE 161
 13 RNPNDKYEPFWEDEKQ-PHMS-----PPELKFQCCGQKTLRPRFKLIGGEFTTIE 64
 162 QPWFPAAYRRHRGGSVTVVCGSLISPCWVISATFCFIDYPKKEDYIVVLRGRSLNSNTQ 221
 65 QPWFPAAYRRHRGGSVTVVCGSLISPCWVISATFCFIDYPKKEDYIVVLRGRSLNSNTQ 124
 222 GEMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQF 281
 125 GEMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQF 184
 282 GTSCEITGFGKENSVDLYPEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPQWKT 341
 185 GTSCEITGFGKENSVDLYPEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPQWKT 244
 342 DSCQDGGGLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGL 401
 245 DSCQDGGGLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGL 304
 402 AL 403
 305 AL 306

RESULT 19
 US-08-944-483-73
 Sequence 73, Application US/08944483
 Patent No. 6232456
 GENERAL INFORMATION:
 APPLICANT: COHEN, MAURICE
 APPLICANT: COLPITTS, TRACEY L.
 APPLICANT: FRIEDMAN, PAULA N.
 APPLICANT: GRANADOS, EDWARD N.
 APPLICANT: KLASS, MICHAEL R.
 APPLICANT: RUSSELL, JOHN C.
 APPLICANT: STEWART, KENT D.
 APPLICANT: STROUPE, STEVEN D.
 TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
 TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
 TITLE OF INVENTION: OF THE PROSTATE
 NUMBER OF SEQUENCES: 76
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Abbott Laboratories
 STREET: 100 Abbott Park Road
 CITY: Abbott Park
 STATE: IL
 COUNTRY: USA
 ZIP: 60064-3500
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/944,483
 FILING DATE:
 CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6232456e
US-08-944-483-73

Query Match 61.2%; Score 1382; DB 3; Length 253;
Best Local Similarity 100.0%; Pred. No. 4.4e-113;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

151 IIGGEFTTIENQPFFAIYRRHGGSVYVCGSLISPCWVISATHCFIDYPKEDYIV 210
1 IIGGEFTTIENQPFFAIYRRHGGSVYVCGSLISPCWVISATHCFIDYPKEDYIV 60
211 LGRSRLNSNTQGMKFEVENLIHLKDYSDATLAHNDIALLKIRSKGRCQAQPSRTIQT 270
61 LGRSRLNSNTQGMKFEVENLIHLKDYSDATLAHNDIALLKIRSKGRCQAQPSRTIQT 120
271 CLPSMYNDPFGTSCETITGFGKENSTDYLYPEQLKMTVVKLIHSHRECCQPHYGSEVTT 330
121 CLPSMYNDPFGTSCETITGFGKENSTDYLYPEQLKMTVVKLIHSHRECCQPHYGSEVTT 180
331 MLCRAADPQWKTDCQGDGSGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTVSHFLP 390
181 MLCRAADPQWKTDCQGDGSGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTVSHFLP 240
391 IRSHTKEENGLAL 403
241 IRSHTKEENGLAL 253

RESULT 20
US-08-560-098A-49
Sequence 49; Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7

FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8900
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 254 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-49

Query Match 60.9%; Score 1374; DB 2; Length 254;
Best Local Similarity 99.2%; Pred. No. 2.2e-112;
Matches 252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

150 KIIGGEFTTIENQPFFAIYRRHGGSVYVCGSLISPCWVISATHCFIDYPKEDYIV 209
1 KIIGGEFTTIENQPFFAIYRRHGGSVYVCGSLISPCWVISATHCFIDYPKEDYIV 60
210 YLGRSRLNSNTQGMKFEVENLIHLKDYSDATLAHNDIALLKIRSKGRCQAQPSRTIQT 269
61 YLGRSRLNSNTQGMKFEVENLIHLKDYSDATLAHNDIALLKIRSKGRCQAQPSRTIQT 120
270 ICLPSMYNDPFGTSCETITGFGKENSTDYLYPEQLKMTVVKLIHSHRECCQPHYGSEVTT 329
121 ICLPSMYNDPFGTSCETITGFGKENSTDYLYPEQLKMTVVKLIHSHRECCQPHYGSEVTT 180
330 KMLCAADPQWKTDCQGDGSGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTVSHFLP 389
181 KMLCAADPQWKTDCQGDGSGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTVSHFLP 240
390 WIRSHTKEENGLAL 403
241 WIRSHTKEENGLAL 254

US-09-101-272G-73
Sequence 73; Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: Q50979
CURRENT APPLICATION NUMBER: US/09/101,272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patent In version 3.1
SEQ ID NO 73
LENGTH: 200
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: ATP domain of uPA
US-09-101-272G-73

Query Match 43.6%; Score 983; DB 4; Length 200;
Best Local Similarity 95.6%; Pred. No. 2.5e-78;
Matches 172; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

1 SNELHQVPSNCDCLNGTGVSNKYFNSNIHWCNCKKFGGCHCEIDKSKTCYEGNGHYRG 60
21 SNELHQVPSNCDCLNGTGVSNKYFNSNIHWCNCKKFGGCHCEIDKSKTCYEGNGHYRG 80

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQVGLK 120
 Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQVGLK 140
 QY 121 PLVQECNVHDCADGK-----LKPCGOKTLRPRFKIIGBEFTTIENOPFAAIYRHH 172
 Df 141 PLVQECNVHDCADGKPPSPBELKPCGOKTLRPRFKIIGBEFTTIENOPFAAIYRHH 200

RE 22
 US 3-560-098A-51
 ; Juence 51, Application US/08560098A
 ; Tent No. 5978841
 ; GENERAL INFORMATION:
 ; APPLICANT: WNENDT, Stephan
 ; APPLICANT: HEINZEL-WIELAND, Regina
 ; APPLICANT: STEFFENS, Gerd Josef
 ; TITLE OF INVENTION: Proteins having Fibrinolytic and
 ; TITLE OF INVENTION: Coagulation-inhibiting Properties
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 ; STREET: 1200 G Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30 (BPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/560,098A
 ; FILING DATE: 17-NOV-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: P 44 40 892.7
 ; FILING DATE: 17-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: EVANS, Joseph D.
 ; REGISTRATION NUMBER: 26,269
 ; REFERENCE/DOCKET NUMBER: 148/42448
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-8800
 ; TELEFAX: (202) 628-8844
 ; INFORMATION FOR SEQ ID NO: 51:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 477 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; 8-560-098A-51
 ;
 ; Query Match 38.4%; Score 867.5; DB 2; Length 477;
 ; st Local Similarity 43.7%; Pred. No. 9.4e-68;
 ; Matches 178; Conservative 60; Mismatches 148; Indels 21; Gaps 9;
 ;
 ; 3 ELHQVP----SNCDLNGGTCVSNKYFSNIHWCNCPKFGGHCIDKSKTCYEGNGHFI 58
 ; 78 QCHTVPVKSCSELRCFNGGTCWQAASFSDF-VQCPCPKGYTGKQCEVDTHATCYKDGVTY 136
 ; 59 RGRKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQV 118
 ; 137 RGTWSTSESAQCLNNMNSLRTARTYNGRSDAITLGLGNHNYCRPNDRNNKFWCVIKA 196
 ; 119 LKPLVQECNVHDCADGKLFQCG-QKTLRPRFKIIGSEFTTIENOPFAAIYRHHRGG- 176
 ; 197 SKFILEPCSPVCS---KATCGLRKYKEPOLHSTGGLFTDITSHPWQAIPAQNRSSG 252
 ; 177 VTYVCGSLSPWCVISATHCIFD-YPKKEDIYVILGRSLNNTQGMKFEVENILHK 235
 ; 253 ERFLCGGILLISSCWLVFAAHCFQERYPPQHLRWV-LGRTRYRKPGBEOTFEVEKCIVHE 311

QY 236 DYSADTLAHNDIALKIRSGRCAQPSRTIOTICLPSMYNDPQFCTSCIEITGFGKENS 295
 Db 312 EFDDDT--YNDIALLOLKSGSPQCAQESDVRAICLPEANLQLPDWTECELSGYGKHS 369
 QY 296 TDYLYPOLKMTVVKLISHRECQOPHYHSEVITVKMLCAADPOWKT-----DSCQDSDG 349
 Db 370 SSFPYSEQLKEGHRVLRPSRCTSKFLFNKTVTKNMLCAGDTRSGEIHFNVDACQDSDG 429
 QY 350 GPLVCSLQGRMTLTGIVSWGRCALCKDKPGVYTRVSHFLFWIRSHTK 396
 Db 430 GPLVCRNDNEMTLGLIISWVGCGEKDIPGVYTKVNYLGNIRDNMR 476

RESULT 23
 US-08-811-949-39
 ; Sequence 39, Application US/08811949
 ; Patent No. 5940533
 ; GENERAL INFORMATION:
 ; APPLICANT: NIWA, MINEO
 ; APPLICANT: SAITO, YOSHIMASA
 ; APPLICANT: SASAKI, HITOSHI
 ; APPLICANT: HAYASHI, MASAKO
 ; APPLICANT: NOTANI, JOUJI
 ; APPLICANT: KOBAYASHI, MASAKAZU
 ; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
 ; NUMBER OF SEQUENCES: 67
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
 ; ADDRESSEE: P.C.
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 ; CITY: ARLINGTON
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/811,949
 ; FILING DATE: 05-MAR-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F.
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 18-966-0
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-413-3000
 ; TELEFAX: 703-413-2220
 ; INFORMATION FOR SEQ ID NO: 39:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 527 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-811-949-39
 ;
 ; Query Match 38.3%; Score 865.5; DB 2; Length 527;
 ; Best Local Similarity 37.7%; Pred. No. 1.6e-67;
 ; Matches 184; Conservative 56; Mismatches 151; Indels 97; Gaps 10;
 ;
 ; 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGHCIDKSKTCYEGNGHFI 58
 ; 42 QCHSVPVKSCSELPFCFNGGTCQQAALYFSDF-VQCCEPGFAGKCECIDTRATCYEDQGISY 100
 ; 59 RGRKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQV 118
 ; 101 RGTWSTAESGAECTNWNSSALAAQFPYSGRRPDAIRLGLGNHNYCRNDRDCKPWCYVKA 160
 ; 119 LKPLVQECNVHDCADG-----LHPLVQECNVHDCADG----- 134

161 GYSSSEFCSTPACSEGNDCYFGNGSAYRGTHTSLTESGASCLPWNMSMILIGKYTAQNP 220
135 -----XKFKQ-----CG-QKTLRPRFKIIGE 155
221 AQAALGLGKHNYCRNPDGAKPWCHVKNRRLTWECYDVPSCSTCGLRQYQSPQFRIGGL 280
156 FTTIENQWFAAIYRRH-RGGSVTVYCGGLSPCWISATHCFIDYPKKEDYIVYLG 214
281 FADIASHPWQAIAFAKHRRSPGERFLCGGLISSCWILSAAHCFQERFPFHLLTVILGT 340
215 RLNSNTQCEMKFEVENILHKDYADTLAHNDIALKIRSKGRCQAQPSRTIOTICLPS 274
341 YRVVPGEEQKFEVEKYIVHKEPDDT--YDNDIALQLKSDSSRCAQESSVVTVCCLPP 398
275 MYNDPQGTSCETITGFKENSTDYLYPEQLKMTVWXLISHRECOQPHYGVSEVTTMCA 334
399 ADQLPDPWTECELSGKGHEALSPFYSERLKEAHLVLPSSRCTSOHLNLRITVDNMLCA 458
335 AD-----POWKT-DSOQDGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFL 388
459 GDRSGGQANLHDACQDGGPLVCLNDGRMTLVGIISWGLCGGQKDPGVYTKVNYL 518
389 PWIRSHTK 396
519 DWIRDNR 526
RESULT 25
US-07-609-510B-16
Sequence 16, Application US/07609510B
Patent No. 5326700
GENERAL INFORMATION:
APPLICANT: Berg et al.
TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue F
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN.
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/609,510B
FILING DATE: 19901106
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 38.1%; Score 859.5; DB 6; Length 527;
Best Local Similarity 37.7%; Pred. No. 5.3e-67;
Matches 184; Conservative 54; Mismatches 153; Indels 97; Gaps 10;
3 ELHQP-SNCD---CLNGTCVSNKYFNSNIHWCNCPKFGQCHCEIDKSKTCYEGNGHY 58
42 QCHSVVPKSCSEPRFCNGTCQALYPSDF-VCOCEGAGKCEIDTRATCYEDQGISY 100
59 RGKASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYQVG 118
101 RGTWSTAESGAECTNWSALAAQKPYSGRRPDAIRLGLGNHNYCRNPDNRDPKWCYVKA 160
119 LKPLVQECVHDCADG-----LKFKQ----- 139
161 GYSSSEFCSTPACSEGNDCYFGNGSAYRGTHTSLTESGASCLPWNMSMILSHQYTAQNP 220
140 -----CG-QKTLRPRFKIIGE 155
221 AQAALGLGKHNYCRNPDGAKPWCHVKNRRLTWECYDVPSCSTCGLRQYQSPQFRIGGL 280
156 FTTIENQWFAAIYRRH-RGGSVTVYCGGLSPCWISATHCFIDYPKKEDYIVYLG 214
281 FADIASHPWQAIAFAKHRRSPGERHLCGGILISSCWILSAAHCFQERFPFHLLTVILGT 340

215 RLNSNTQCEMKFEVENILHKDYADTLAHNDIALKIRSKGRCQAQPSRTIOTICLPS 274
341 YRVVPGEEQKFEVEKYIVHKEPDDT--YDNDIALQLKSDSSRCAQESSVVTVCCLPP 398
275 MYNDPQGTSCETITGFKENSTDYLYPEQLKMTVWXLISHRECOQPHYGVSEVTTMCA 334
399 ADQLPDPWTECELSGKGHEALSPFYSERLKEAHLVLPSSRCTSOHLNLRITVDNMLCA 458
335 AD-----POWKT-DSOQDGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFL 388
459 GDRSGGQANLHDACQDGGPLVCLNDGRMTLVGIISWGLCGGQKDPGVYTKVNYL 518
389 PWIRSHTK 396
519 DWIRDNR 526
RESULT 25
US-07-609-510B-16
Sequence 16, Application US/07609510B
Patent No. 5326700
GENERAL INFORMATION:
APPLICANT: Berg et al.
TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue F
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN.
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/609,510B
FILING DATE: 19901106
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 38.0%; Score 858.5; DB 1; Length 527;
Best Local Similarity 37.5%; Pred. No. 6.5e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
3 ELHQP-SNCD---CLNGTCVSNKYFNSNIHWCNCPKFGQCHCEIDKSKTCYEGNGHY 58
42 QCHSVVPKSCSEPRFCNGTCQALYPSDF-VCOCEGAGKCEIDTRATCYEDQGISY 100
59 RGKASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYQVG 118
101 RGTWSTAESGAECTNWSALAAQKPYSGRRPDAIRLGLGNHNYCRNPDNRDPKWCYVKA 160
119 LKPLVQECVHDCADG-----LKFKQ----- 134
161 GYSSSEFCSTPACSEGNDCYFGNGSAYRGTHTSLTESGASCLPWNMSMILIGKYTAQNP 220
135 -----CG-QKTLRPRFKIIGE 155
221 AQAALGLGKHNYCRNPDGAKPWCHVKNRRLTWECYDVPSCSTCGLRQYQSPQFRIGGL 280
156 FTTIENQWFAAIYRRH-RGGSVTVYCGGLSPCWISATHCFIDYPKKEDYIVYLG 214
281 FADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPFHLLTVILGT 340

QY 215 RLNSTQEMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCAQPSRTIOTICLPS 274
DE 341 YRVVGEBOQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDSSRCACQSSVVRTVCLPP 398
QY 275 MYNDPQFGTSCHEITGFCOKENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCA 334
DE 399 ADLQLPDWTCELSGYKGHEALSPFYSERLKEARVRLYPSSRCTSQHLLNRTVTDNMLCA 458
QY 335 AD---POWKT--DSCOGDSGGPLVCSLQGRMTLTGIVSMGRGCAKDKPGVTVRYSHFL 388
DE 459 GDTSGGPPQANLHDACQDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVTVKVTNYL 518
QY 389 PWIRSHTK 396
DE 519 DWIRDNMR 526
RE LT 26
PC US91-01025A-2
; quence 2, Application PC/TUS9101025A
; ENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01025A
; FILING DATE: 19910214
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/486,657
; FILING DATE: 1 March 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 454P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1896
; TELEFAX: 415/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PC US91-01025A-2

Db 161 GKYSEFCSTPACSEGNDCYFGNGSAVRGTHSLTESGASCLPWNSMLIGKVYTAQNPS 220
QY 135 -----KLKFO-----CG-OKTLPKRIIGGE 155
Db 221 AQAALGLGKHNYCRNPDGDAKFWCHLVKNNRLTWEYCDVPSCSTGLRQYQSOPFRKIGGL 280
QY 156 FTTIENQPFWFAAIYRRH-RGGSVTYVCGGSLISPCWISATHCFIDYPKKEDIYVILGRS 214
Db 281 FADIASHPWQAARIAKAPRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRT 340
QY 215 RLNSTQEMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCAQPSRTIOTICLPS 274
Db 341 YRVVGEBOQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDSSRCACQSSVVRTVCLPP 398
QY 275 MYNDPQFGTSCHEITGFCOKENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCA 334
Db 399 ADLQLPDWTCELSGYKGHEALSPFYSERLKEARVRLYPSSRCTSQHLLNRTVTDNMLCA 458
QY 335 AD---POWKT--DSCOGDSGGPLVCSLQGRMTLTGIVSMGRGCAKDKPGVTVRYSHFL 388
Db 459 GDTSGGPPQANLHDACQDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVTVKVTNYL 518
QY 389 PWIRSHTK 396
Db 519 DWIRDNMR 526
RESULT 27
5185259-8
; Patent No. 5185259
; APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHAR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO:8:
; LENGTH: 527
5185259-8
Query Match 38.0%; Score 858.5; DB 5; Length 527;
Best Local Similarity 37.5%; Pred. No.6.5e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFNSIHWNCNPKFGQHCHEIDKSKTCYEGNGHFY 58
Db 42 QCHSVFVKSCSEPRCFNGGTCCQALYFSDF-VCOCPGFGAGKCCCEIDTRATCYEDQGISY 100
QY 59 RGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
Db 101 RGTWSTAGSAGACTNWSNLSAQAQPYSGRRPDAIRLGLGNHNYCRNPDSDKPCWYFKA 160
QY 119 LKPLVQECMVHDCADG-----KLKFO-----CG-OKTLPKRIIGGE 155
Db 161 GKYSEFCSTPACSEGNDCYFGNGSAVRGTHSLTESGASCLPWNSMLIGKVYTAQNPS 220
QY 135 -----KLKFO-----CG-OKTLPKRIIGGE 155
Db 221 AQAALGLGKHNYCRNPDGDAKFWCHLVKNNRLTWEYCDVPSCSTGLRQYQSOPFRKIGGL 280
QY 156 FTTIENQPFWFAAIYRRH-RGGSVTYVCGGSLISPCWISATHCFIDYPKKEDIYVILGRS 214

I 281 FADIASHPWQAALFAKRRSPGFRFLCGGILSSCWILSAHCFQERFFPHLTVILGRT 340
C 215 RLSNTQGMKFEVENILHKOVSADTLAHNDIALKIRSKRGCAQPSRTIQTICLPS 274
I 341 YRVVPEEEQKFEVEKYIVHKEPDDDT--YDNDIALQLKSDSRCAQESSVVRTVCLPP 398
C 275 MYNDPQFGTSCETITGKENSVDLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCA 334
I 399 ADLQLPDWTCELSGYKGHEALSPFYSERLKEAHVLYPSSRCTSOHLNRTVTDNMLCA 458
C 335 AD-----PWKT-DSQCGSDGGLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFL 388
I 459 GDRSGGPQANLHDACQSDGGLVCLNDGRMTLVGIISWGLCGQKDVGVYTKVTNYL 518
C 389 PWIRSHTK 396
I 519 DWIRDNR 526

I 08-811-949-43
sequence 43, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SAKAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
08-811-949-43

very Match 38.0%; Score 858.5; DB 2; Length 562;
est Local Similarity 37.5%; Pred. No. 7.1e-67;
atches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
3 ELHQVP-SNCD---CLNGGTVCVNKYFSNIHNCNCPKFTGGQHCEDKSKTCYEGNGHFY 58
77 QCHSVPVKSCSEPCFNGGTCCQALYFSDP-VQCPEGFAGKCCEDTTRATCYEQGISY 135

QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHSDALQLGLGKHNYCRNPDNRERPKCYQVG 118
DB 136 RGTWSTAESGAECNTWNSALLAQKPYSGRRPDATIRLGLGNHNYCRNPDNRDPSKPYVFK 195
QY 119 LKPJVQECMVHDCADG----- 134
DB 196 KYSSEFCSTPACSEGNDCVFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPS 255
QY 135 -----KLKFO-----CG-OKTLPKPKIIGGE 155
DB 256 AQAALGLGKHNYCRNPDGAKPWCHLVKNRRLTWEDVPCSTCGLRQYQOPQRIKGL 315
QY 156 FTTIENQWFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRS 214
DB 316 FADIASHPWQAALFAKRRSPGFRFLCGGILSSCWILSAHCFQERFFPHLTVILGRT 375
QY 215 RLSNTQGMKFEVENILHKOVSADTLAHNDIALKIRSKRGCAQPSRTIQTICLPS 274
DB 376 YRVVPEEEQKFEVEKYIVHKEPDDDT--YDNDIALQLKSDSRCAQESSVVRTVCLPP 433
QY 275 MYNDPQFGTSCETITGKENSVDLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCA 334
DB 434 ADLQLPDWTCELSGYKGHEALSPFYSERLKEAHVLYPSSRCTSOHLNRTVTDNMLCA 493
QY 335 AD-----PWKT-DSQCGSDGGLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFL 388
DB 494 GDRSGGPQANLHDACQSDGGLVCLNDGRMTLVGIISWGLCGQKDVGVYTKVTNYL 553
QY 389 PWIRSHTK 396
DB 554 DWIRDNR 561

RESULT 29
US-08-560-098A-50
Sequence 50, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WNEENDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evensen, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US: 8-560-098A-50

ery Match
; 38.0%; Score 858.5; DB 2; Length 562;
; st Local Similarity 37.5%; Pred. No. 7.1e-67;
; tches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

Q: 3 ELHQP--SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQCEIDKSKTKCYEGNGHFY 58
D: 77 QCHSVPVKSCSEPCFNGGTCQALYFSDP--VCQPEGFAGKCEIDTRATCYEDQGISY 135
Q: 59 RGKASTDTMGPRCLPWNASATVLOQTYHAHRSDALOLGLGKHNCRNPDNRPRPCVQVG 118
D: 136 RGTWSTABSGAECTNWNSSALAAQKPYSGRRPDALRLGLGNHNYCRNPDNRDRSKPCVYFKA 195
Q: 119 LKPLVQECMVHDCADG----- 134
D: 196 KYSSEFCSTPACSEGNSDCYFGNGSAVRGTHSITSGASCLPWNMILIGKYVTAQNP 255
Q: 135 -----KLKFO----- 155
D: 256 AQAALGLGKHNCRNPDGDAKPWCHLVKNRLTWECVDPSCSTCGLRQYSQPFRIKGG 315
Q: 156 FTTIENQWFAAIYRRH--RGGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVVLGRS 214
D: 316 PADIASHPWQAIAFAKRRSPGERFLCGGILISSCWILSAHCFQERFPFPHLTVILGR 375
Q: 215 RLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSKGRCACQPSRTIOTICLPS 274
D: 376 YRVVPGEEQKFEVEKIVHKEFDDT--YDNDIALQLKSDSRCAQESSVVRTVCLPP 433
Q: 275 MYNDPQFCTSCITGFGKSTDLVYPEQLKMTVVKLISHRECQCPHYGSEVITKMLCA 334
D: 434 ADLQLPDWTCELSGKGKHEALSPYSERLKEAHVRLYPSRCTSQHLLNRTVDNMLCA 493
Q: 335 AD-----PWKMT--DSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 388
D: 494 GDTSGGPOANLHDACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDPVGYTKVTNYL 553
Q: 389 PWIRSHTK 396
D: 554 DWIRDNR 561

; RESULT 31
; US-09-703-695A-4
; Sequence 4, Application US/09703695A
; Patent No. 6593097
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: P1788R1
; CURRENT APPLICATION NUMBER: US/09/703,695A
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens

```

09-703-695A-4
Query Match 38.0%; Score 858.5; DB 4; Length 562;
est Local Similarity 37.5%; Pred. No. 7.1e-67;
atches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
3 ELHQP-SNCD---CLNGGTCVSNKYPNSATVLOQTYHAHSDALQGLGKHNYCRNPNRRPWCYQVG 58
77 QCHSVPVKSCSEPRCFNGGTCQOALYFSDP-VQCPGEPGAGKCEIDTRATCYEDQGISY 135
59 RKGASTDTMGRPCLPWNSATVLOQTYHAHSDALQGLGKHNYCRNPNRRPWCYQVG 118
136 RGTWSTAESGAECTNWNSSALAKPYSGRPPDAIRLGLGNHNYCRNPNDRSKPWCYVFKA 195
119 LKPLVQECMVHDCADG----- 134
196 KYSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNP 255
135 -----KLKQ-----CG-OKTLRPREKIIIGE 155
256 AQAALGLGKHNYCRNPDGDAKPWCHLVKNRRLTWECYDVPSCSTGLRQYQOPRIKGG 315
156 FTTIENQFWFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIYVLG 214
316 FADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILG 375
215 RLNSNTQGENKFEVENILHKOYSADTLAHHNDIALKIRSKBGRCAQPSRTIQTICLPS 274
376 YRVVPGEEOKFEVEKYIVHKEPDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 433
275 MYNDPQGTSCETITGFGKENSTDYLPQOLKMTVVKLIHSHRECOQPHYVGSVTTMCA 334
434 ADLQPDWTECELSGKGHEALSPFYSERLKEAHVLYPSSRCTSQHLLNRVTDNMLCA 493
335 AD-----PQWKT-DSQCGSGGGLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFL 388
494 GDRSGGPQANLHDACQSGGGLVCLNDGRMTLVGIISWGLCGGQKQVPGVYTKVTNYL 553
389 PWIRSHTK 396
554 DWIRDNR 561
RESULT 33
5200340-2
; Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,
; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
; ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
; SEQ ID NO:2:
; LENGTH: 562
5200340-2
Query Match 38.0%; Score 858.5; DB 6; Length 562;
est Local Similarity 37.5%; Pred. No. 7.1e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
3 ELHQP-SNCD---CLNGGTCVSNKYPNSATVLOQTYHAHSDALQGLGKHNYCRNPNRRPWCYQVG 58
77 QCHSVPVKSCSEPRCFNGGTCQOALYFSDP-VQCPGEPGAGKCEIDTRATCYEDQGISY 135
59 RKGASTDTMGRPCLPWNSATVLOQTYHAHSDALQGLGKHNYCRNPNRRPWCYQVG 118
136 RGTWSTAESGAECTNWNSSALAKPYSGRPPDAIRLGLGNHNYCRNPNDRSKPWCYVFKA 195
119 LKPLVQECMVHDCADG----- 134
196 KYSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNP 255
135 -----KLKQ-----CG-OKTLRPREKIIIGE 155
256 AQAALGLGKHNYCRNPDGDAKPWCHLVKNRRLTWECYDVPSCSTGLRQYQOPRIKGG 315
156 FTTIENQFWFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIYVLG 214
316 FADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILG 375
215 RLNSNTQGENKFEVENILHKOYSADTLAHHNDIALKIRSKBGRCAQPSRTIQTICLPS 274
376 YRVVPGEEOKFEVEKYIVHKEPDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 433
275 MYNDPQGTSCETITGFGKENSTDYLPQOLKMTVVKLIHSHRECOQPHYVGSVTTMCA 334
434 ADLQPDWTECELSGKGHEALSPFYSERLKEAHVLYPSSRCTSQHLLNRVTDNMLCA 493
335 AD-----PQWKT-DSQCGSGGGLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFL 388
494 GDRSGGPQANLHDACQSGGGLVCLNDGRMTLVGIISWGLCGGQKQVPGVYTKVTNYL 553
389 PWIRSHTK 396
554 DWIRDNR 561
ULT 32
5259-3
; Patent No. 5185259
; APPLICANT: GOEDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; HAR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; Q ID NO:3:
; LENGTH: 562
5259-3
Query Match 38.0%; Score 858.5; DB 6; Length 562;
est Local Similarity 37.5%; Pred. No. 7.1e-67;
atches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
3 ELHQP-SNCD---CLNGGTCVSNKYPNSATVLOQTYHAHSDALQGLGKHNYCRNPNRRPWCYQVG 58

Db 77 QCHSVPVKSCSEPRCFNGGTCQOALYFSDP-VQCPGEPGAGKCEIDTRATCYEDQGISY 135
Qy 59 RKGASTDTMGRPCLPWNSATVLOQTYHAHSDALQGLGKHNYCRNPNRRPWCYQVG 118
Db 136 RGTWSTAESGAECTNWNSSALAKPYSGRPPDAIRLGLGNHNYCRNPNDRSKPWCYVFKA 195
Qy 119 LKPLVQECMVHDCADG----- 134
Db 196 KYSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNP 255
Qy 135 -----KLKQ-----CG-OKTLRPREKIIIGE 155
Db 256 AQAALGLGKHNYCRNPDGDAKPWCHLVKNRRLTWECYDVPSCSTGLRQYQOPRIKGG 315
Qy 156 FTTIENQFWFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIYVLG 214
Db 316 FADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILG 375
Qy 215 RLNSNTQGENKFEVENILHKOYSADTLAHHNDIALKIRSKBGRCAQPSRTIQTICLPS 274
Db 376 YRVVPGEEOKFEVEKYIVHKEPDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 433
Qy 275 MYNDPQGTSCETITGFGKENSTDYLPQOLKMTVVKLIHSHRECOQPHYVGSVTTMCA 334
Db 434 ADLQPDWTECELSGKGHEALSPFYSERLKEAHVLYPSSRCTSQHLLNRVTDNMLCA 493
Qy 335 AD-----PQWKT-DSQCGSGGGLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFL 388
Db 494 GDRSGGPQANLHDACQSGGGLVCLNDGRMTLVGIISWGLCGGQKQVPGVYTKVTNYL 553
Qy 389 PWIRSHTK 396
Db 554 DWIRDNR 561
RESULT 33
5200340-2
; Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,
; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
; ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
; SEQ ID NO:2:
; LENGTH: 562
5200340-2
Query Match 38.0%; Score 858.5; DB 6; Length 562;
est Local Similarity 37.5%; Pred. No. 7.1e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
3 ELHQP-SNCD---CLNGGTCVSNKYPNSATVLOQTYHAHSDALQGLGKHNYCRNPNRRPWCYQVG 58
77 QCHSVPVKSCSEPRCFNGGTCQOALYFSDP-VQCPGEPGAGKCEIDTRATCYEDQGISY 135
59 RKGASTDTMGRPCLPWNSATVLOQTYHAHSDALQGLGKHNYCRNPNRRPWCYQVG 118
136 RGTWSTAESGAECTNWNSSALAKPYSGRPPDAIRLGLGNHNYCRNPNDRSKPWCYVFKA 195
119 LKPLVQECMVHDCADG----- 134
196 KYSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNP 255
135 -----KLKQ-----CG-OKTLRPREKIIIGE 155
256 AQAALGLGKHNYCRNPDGDAKPWCHLVKNRRLTWECYDVPSCSTGLRQYQOPRIKGG 315
156 FTTIENQFWFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIYVLG 214
316 FADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILG 375
215 RLNSNTQGENKFEVENILHKOYSADTLAHHNDIALKIRSKBGRCAQPSRTIQTICLPS 274
376 YRVVPGEEOKFEVEKYIVHKEPDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 433
275 MYNDPQGTSCETITGFGKENSTDYLPQOLKMTVVKLIHSHRECOQPHYVGSVTTMCA 334
434 ADLQPDWTECELSGKGHEALSPFYSERLKEAHVLYPSSRCTSQHLLNRVTDNMLCA 493
335 AD-----PQWKT-DSQCGSGGGLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFL 388
494 GDRSGGPQANLHDACQSGGGLVCLNDGRMTLVGIISWGLCGGQKQVPGVYTKVTNYL 553
389 PWIRSHTK 396
554 DWIRDNR 561

QY 215 RLNSNTGEMKFEVENILHKOYSADTLAHNDIALKIRSEKGRCAQPSRTIOTICLPS 274
 DE 376 YRVVPGEEQKFEVEKIYHKEFDDT--YNDIALQLKSDSSRCAQESSVVRVTCVCLPP 433
 QY 275 MYNDPQFTSCITGFKENSTDYLYPEQLKMTVVKLISHRECCQPHYGSEVTTMCLCA 334
 DE 434 ADLQLPDWTCELSGYGKHEALSPFYSERLKEAHRVLPSSRCTSQHLLNRTVTDNMLCA 493
 QY 335 AD-----POWKT-DSCOGDSGGLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFL 388
 DE 494 GDRSGGQANLHDACCQDSGGLVCLNDGRMTLVGLISWGLGCGQKDPGVYTKVTNYL 553
 QY 389 PWIRSHTK 396
 DE 554 DWIRDNR 561
 RE LT 34
 53 773-2
 ; Applicant: WEI, CHA-MER;HSIUNG, NANCY;REDDY, VERMURI B.;
 ; ONT, JEFFREY F.;DACKOWSKI, WILLIAM;DOUGLAS, RICHARD;
 ; E, EDWARD S.;PURCELL JR., RICHARD D.;LAU, DAVID TAI-YUI
 ; TITLE OF INVENTION: HUMAN UTERINE TISSUE PLASMINOGEN
 ; IVATOR PRODUCED BY RECOMBIANT DNA
 ; NUMBER OF SEQUENCES: 6
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/782,686
 ; FILING DATE: 01-OCT-1985
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 656,770
 ; FILING DATE: 01-OCT-1984
 ; ID NO.2:
 ; LENGTH: 562
 ; 773-2

ary Match 38.0%; Score 858.5; DB 6; Length 562;
 st Local Similarity 37.5%; Pred. No. 7.1e-67;
 tches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
 3 ELHVP-SNCD---CLNGGTCVSNKYPSNTHWNCNPKKFGQHCEDKSKTCEGNGHGY 58
 77 QCHSVPVKSCSEPRCFNGGTCQQALYPSDF-VCOCPGFGAGKCEIDTRATCYEDQGISY 135
 59 RGKASTDTMGPRCLPWNASATVLOQTYHAHRSADALQGLGKHNYCRNPNRPPWCYVQVG 118
 136 RGTWSTAESGAECTNWNSSALAKQPYSGRRPDAILRGLGNHNYCRNPDSDSKPCVYFKA 195
 119 LKPLVQECMVHDCADG----- 134
 196 KYSGSEFCSTPACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNP 255
 135 -----KLKFO-----CG-OKTLRPRFKIIGGE 155
 256 AQALGLGKHNYCRNPDGAKPWCHLVKNRLTWECVDPSCSTCGLRQYSPQFRKGG 315
 156 PTTIENQFWFAIYVRH-RGGSVTVCGSLSPCWVISATHCFIDYPKKEDYIVYLGSR 214
 316 FADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSSAAHCFQERFPHPHLLTVILGRT 375
 215 RLNSNTGEMKFEVENILHKOYSADTLAHNDIALKIRSEKGRCAQPSRTIOTICLPS 274
 376 YRVVPGEEQKFEVEKIYHKEFDDT--YNDIALQLKSDSSRCAQESSVVRVTCVCLPP 433
 275 MYNDPQFTSCITGFKENSTDYLYPEQLKMTVVKLISHRECCQPHYGSEVTTMCLCA 334
 434 ADLQLPDWTCELSGYGKHEALSPFYSERLKEAHRVLPSSRCTSQHLLNRTVTDNMLCA 493
 335 AD-----POWKT-DSCOGDSGGLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFL 388
 494 GDRSGGQANLHDACCQDSGGLVCLNDGRMTLVGLISWGLGCGQKDPGVYTKVTNYL 553

QY 389 PWIRSHTK 396
 DE 554 DWIRDNR 561
 RESULT 35
 5244676-5
 ; Patent No. 5244676
 ; APPLICANT: BELL, LESLIE D.;MAVER, ERNEST J.;PALMIER, MARK O.
 ; TOLUNAY, H.ESER;WARREN, THOMAS G.;WUN, TZE-CHIN
 ; TITLE OF INVENTION: MODIFIED TISSUE PLASMINOGEN ACTIVATOR
 ; WITH MODIFIED GLYCOSYLATION SITE
 ; NUMBER OF SEQUENCES: 9
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/203,047
 ; FILING DATE: 06-JUN-1988
 ; SEQ ID NO.5:
 ; LENGTH: 562
 ; 5244676-5

Query Match 37.7%; Score 851.5; DB 6; Length 562;
 Best Local Similarity 37.1%; Pred. No. 2.9e-66;
 Matches 181; Conservative 58; Mismatches 152; Indels 97; Gaps 10;
 3 ELHVP-SNCD---CLNGGTCVSNKYPSNTHWNCNPKKFGQHCEDKSKTCEGNGHGY 58
 77 QCHSVPVKSCSEPRCFNGGTCQQALYPSDF-VCOCPGFGAGKCEIDGNSDCYPGSGAY 135
 59 RGKASTDTMGPRCLPWNASATVLOQTYHAHRSADALQGLGKHNYCRNPNRPPWCYVQVG 118
 136 RGTWSTAESGAECTNWNSSALAKQPYSGRRPDAILRGLGNHNYCRNPDSDSKPCVYFKA 195
 119 LKPLVQECMVHDCADG----- 134
 196 RRLTWECVDPSCSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNP 255
 135 -----KLKFO-----CG-OKTLRPRFKIIGGE 155
 256 AQALGLGKHNYCRNPDGAKPWCHLVKNRLTWECVDPSCSTCGLRQYSPQFRKGG 315
 156 PTTIENQFWFAIYVRH-RGGSVTVCGSLSPCWVISATHCFIDYPKKEDYIVYLGSR 214
 316 FADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSSAAHCFQERFPHPHLLTVILGRT 375
 215 RLNSNTGEMKFEVENILHKOYSADTLAHNDIALKIRSEKGRCAQPSRTIOTICLPS 274
 376 YRVVPGEEQKFEVEKIYHKEFDDT--YNDIALQLKSDSSRCAQESSVVRVTCVCLPP 433
 275 MYNDPQFTSCITGFKENSTDYLYPEQLKMTVVKLISHRECCQPHYGSEVTTMCLCA 334
 434 ADLQLPDWTCELSGYGKHEALSPFYSERLKEAHRVLPSSRCTSQHLLNRTVTDNMLCA 493
 335 AD-----POWKT-DSCOGDSGGLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFL 388
 494 GDRSGGQANLHDACCQDSGGLVCLNDGRMTLVGLISWGLGCGQKDPGVYTKVTNYL 553
 QY 389 PWIRSHTK 396
 DE 554 DWIRDNR 561

RESULT 36
 US-08-142-590B-25
 ; Sequence 25, Application US/08142590B
 ; Patent No. 6120765
 ; GENERAL INFORMATION:
 ; APPLICANT: HIBINO, Tashihiko; TAKAHASHI, Tadahito; HORII, Izumi; and
 ; TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD, LLP
 ; STREET: 28 State Street
 ; CITY: Boston

```
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,590B
FILING DATE: 25-OCT-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/042,318
FILING DATE: 02-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-009CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-142-590B-25

Query Match 36.6%; Score 825; DB 3; Length 157;
Best Local Similarity 91.7%; Pred. No. 1.2e-64;
Matches 144; Conservative 0; Mismatches 5; Indels 8; Gaps 1;

1 SNEHLQVPSNCDCLNGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
1 SNEHLQVPSNCDCLNGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
61 KASDTMTGRCLPWSATVLOQTYHAHRSALQGLGKHYCNPNRRPRPCYVQVGLK 120
61 KASDTMTGRCLPWSATVLOQTYHAHRSALQGLGKHYCNPNRRPRPCYVQVGLK 120
121 PLVQECMVHDCADGK-----LKFQCGQKTLRPRF 149
121 PLVQECMVHDCADGKSPPELKFQCGQKYLPRF 157

ULT 37
Sequence 59, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-59

Query Match 35.8%; Score 808; DB 2; Length 355;
Best Local Similarity 45.9%; Pred. No. 1e-62;
Matches 163; Conservative 44; Mismatches 132; Indels 16; Gaps 6;

QY 50 CYEGNGHYRGKASDTMTGRCLPWSATVLOQTYHAHRSALQGLGKHYCNPNRR 109
DB 8 CYFGNGSAYRGTHSLITESGASCLPWSNMLIGKVYTAQNPSAALGLGKHYCNPDGA 67
QY 110 RPWCYVQGLKPLVQECMVHDCADGKLFQCG-----TGCLQTLRPRFKIIGGFTTIENOPWFAAI 168
DB 68 KPWCVLKVRRLTWEYCDVPCS-----TCGLQTLRPRFKIIGGFTTIENOPWFAAI 121
QY 169 YRRH-RGGSVTVYCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFE 227
DB 122 FAKHRRSPGERFLCGGILISSCWLSAAHCFQERFPFPHLTVILGRTVYVVGEEQKFE 181
QY 228 VENILKHDYSADTLAHNDIALKIRSKGRCAPSTTQICLPSMYNDPQFGTSCEI 287
DB 182 VEKIVHKEFDDT--YDNDIALQLKSDSRCAQSSVVRVTVCLPADLQLPDWTCECL 239
QY 288 TGFGENSTDYLYPEQLKMTVVKLIHRECOQPHYGYSEVTTKMLCAAD-----PWKRT- 341
DB 240 SGYKHEALSFPYSERLKEAHVRLYPSRCTSQHLLNRTVTDNMLCAGDTRSGGQANLH 299
QY 342 DSCGDSGGPLVCSLQGRMTLTGIVSGRGALCKDFGVTRVSHFLPWIRSHTK 396
DB 300 DACQDSGGPLVCLNDGRMTLVGIISWGLCGQKDVFPVYTKVTNYLDWIRDNR 354

RESULT 38
US-08-811-949-51
Sequence 51, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
```


LENGTH: 208
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: ATPI-ML chimeric protein
09-101-2726-98

very Match 35.4%; Score 799; DB 4; Length 208;
est Local Similarity 94.6%; Pred. No. 3.2e-62;
atches 139; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

1 SNEHVPNCCLANGCTCVSNKYFSNIHWCNPKKFGGHCIDKSKTCYEGNGHFYRG 60
2 SNEHVPNCCLANGCTCVSNKYFSNIHWCNPKKFGGHCIDKSKTCYEGNGHFYRG 61
61 KASTDTMGPRCLPWSATVLOQTYHAHRSADALQGLGKHYCRNPDNRERRPWCYVQVGLX 120
62 KASTDTMGPRCLPWSATVLOQTYHAHRSADALQGLGKHYCRNPDNRERRPWCYVQVGLX 121
121 PLVQECMVHDCADGK-----LKQ 139
122 PLVQECMVHDCADGKPPSPBELKQ 148

ULT 41
08-811-949-63
quence 63, Application US/08811949
tent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
08-811-949-63

very Match 35.3%; Score 796; DB 2; Length 472;
est Local Similarity 42.8%; Pred. No. 1.7e-61;
atches 163; Conservative 51; Mismatches 145; Indels 22; Gaps 8;

30 WCNCPK--KFGQHCEI----DKSKTCYEGNGHFYRGKASTDTMGPRCLPWSATVLOQT 83
99 WCVVFKAGKYSBFCSTPACSEGNSDCYFGNGSAYRGTHSLTBSGASCLPWNMILIGV 158
84 YHAHRSADALQGLGKHYCRNPDNRERRPWCYVQVGLKPLVQECMVHDCADGKPKFGG-Q 142
159 YTAQNPQAALGLGKHYCRNPDGDAKPMCHVLRNRLTWECYCDVPSCS-----TCGLR 212
143 KTLRPRFKITGGEFTTIENQWFAAIYRRH-RGGSVTVYVCGGSLISPCWVISATHCFIDY 201
213 QYQOPQRIKGLPADIASHPWQAIAFAKHRRSPGFRFLCGGILISCCWILSAKCFQER 272
202 PKEDYIVYLGRLNSNTQGENKFEVENILHKOYSADTLAHNDIALKIRKSGRCA 261
273 FPHHLTVILGRTRYRVVPGEEQKFEVKYIVHKFDDDT--YNDIALIQLKSDSRCA 330
262 QPRTIOTICLPMSYNDPQGTSCETITFGKENSTDYLYPEOLKMTWKLISHRECOQPH 321
331 QESVVRVTVCLPADLQLPDWTCELSGYGKHEALSPFYSERLKEAHRVLYPSRCTSQH 390
322 YGSEVTTKMLCAAD-----PQWKT-DSQGDGSGPLVCSLQGRMTLTGIVSWGRGALK 375
391 LLNRVTVDNMLCAGDTRSGGPQANLHDACQDQSGGPLVCLINDGRMTLVGLIISWGLGCGQK 450
376 DKPGVTVRVSHFLPWIRSHTK 396
451 DVGVTYKVTNYLDWINDNR 471

RESULT 42
US-08-797-689-12
Sequence 12, Application US/08797689
Patent No. 587969
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
APPLICANT: Fournier, Alain
APPLICANT: Guittou, Jean-Dominique
APPLICANT: Jung, Gerard
APPLICANT: Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839

```

TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-880-503-6-rai

Query Match 35.1%; Score 793; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 6.3e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 4 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 63
QY 61 KASDTMTGRPCLPWNSATVLTQYTHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 64 KASDTMTGRPCLPWNSATVLTQYTHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 44
US-08-811-949-55
Sequence 55, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-55

Query Match 35.1%; Score 793; DB 2; Length 437;
Best Local Similarity 42.8%; Pred. No. 2.7e-61;
Matches 163; Conservative 50; Mismatches 146; Indels 22; Gaps 8;

```

```

TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-880-503-6-rai

Query Match 35.1%; Score 793; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 6.3e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 4 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 63
QY 61 KASDTMTGRPCLPWNSATVLTQYTHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 64 KASDTMTGRPCLPWNSATVLTQYTHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 44
US-08-811-949-55
Sequence 55, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-55

Query Match 35.1%; Score 793; DB 2; Length 437;
Best Local Similarity 42.8%; Pred. No. 2.7e-61;
Matches 163; Conservative 50; Mismatches 146; Indels 22; Gaps 8;

```


SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
8-411-949-47

Query Match 35.0%; Score 791; DB 2; Length 355;
st Local Similarity 44.8%; Pred. No. 3.1e-61;
Matches 159; Conservative 49; Mismatches 133; Indels 16; Gaps 6;
QY 50 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPDR 109
DB 8 CYFNGSAYGTHSLTESGASCLPWNMILIGVYTAQNPASQALGLGKHNYCRNPDDGA 67
QY 110 RPKYVQVGLKPLVQECMVHDCADGLKXFCQG-QKTLRPRFKIIGSEFTTIENQWFAAI 168
DB 68 KPWCHLKNRLTWECVDPSCS-----TCGLRQYSQPQFRIIGLGFADIASHPQAAI 121
QY 169 YRSH-RGGSVTVYVCGSLISPCWVISAHCDFIDYPKKEDYIVYLGSRSLNSNTQEMKPE 227
DB 122 FAKHRSPPGRFUCGGILISSCWILSAHCFQERPPPHLTVILGRTYRVVPGEEQKPE 181
QY 228 VENLIHKDYSADTLAHNDIALLKIRSKESGCAQPSRTIQTICLPSMYNDPQFGTSCBI 287
DB 182 VEKVIHKEFDDT--YNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCEEL 239
QY 288 TFGCKENSTLYLPQOLKMTVVKLISHRECQCPHYGSEVTTKMLCAAD-----PWKT- 341
DB 240 SGYKGHEALSPFYSERLKEAHVRLYPSRSTSOHLNRTVTDNMLCAGDTRSGGPQANLH 299
QY 342 DSCQGSQGLPVCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTK 396
DB 300 DACQGSQGLPVCCLNDGRMTLVGLISWGLGCGQKDPGVYTKVNYLDWIRDNR 354

BT 47
US 8-427-640-8
; Invention 8, Application US/08427640
; Patent No. 5658788
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Compounds and Methods for Treatment of Thromboembolic Disor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,640
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/689,410
; FILING DATE: 22 APRIL 1991
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 356 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; 8-427-640-8
; Query Match 35.0%; Score 790; DB 1; Length 356;
; st Local Similarity 44.2%; Pred. No. 3.9e-61;

Matches 159; Conservative 49; Mismatches 136; Indels 16; Gaps 6;
QY 45 DKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRN 104
DB 4 DTRATCYEDQGISYRGTWSTAEGACTNWNSSALAKQPSGRRPPDAIRLGLGHNHNYCRN 63
QY 105 PDNRPRWCYVQVGLKPLVQECMVHDCADGLKXFCQG-QKTLRPRFKIIGSEFTTIENQ 163
DB 64 PDGDAKWCCHLVKORRLTWECVDPSCS-----TCGLRQYSQPQFRIIGLGFADIASHP 117
QY 164 WFAAIYRSH-RGGSVTVYVCGSLISPCWVISAHCDFIDYPKKEDYIVYLGSRSLNSNTQ 222
DB 118 WQAAIFAKHRSPPGRFELCGGILISSCWILSAHCFQERPPPHLTVILGRTYRVVPGEE 177
QY 223 EMKEFEVENLIHKDYSADTLAHNDIALLKIRSKESGCAQPSRTIQTICLPSMYNDPQFG 282
DB 178 EOKFEVEKYIVHKEFDDT--YNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDW 235
QY 283 TSCITFGCKENSTLYLPQOLKMTVVKLISHRECQCPHYGSEVTTKMLCAAD-----P 337
DB 236 TECELSGYKGHEALSPFYSERLKEAHVRLYPSRSTSOHLNRTVTDNMLCAGDTRSGGP 295
QY 338 QWKT-DSCQGSQGLPVCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTK 396
DB 296 QANLHDACQGSQGLPVCCLNDGRMTLVGLISWGLGCGQKDPGVYTKVNYLDWIRDNR 355

RESULT 48
US-09-101-272G-80
; Sequence 80, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: QS0979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 80
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFHI chimeric protein
US-09-101-272G-80

Query Match 34.9%; Score 788; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 2.7e-61;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSHNWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 2 SNELHQVPSNCDCLNGGTCVSNKYFNSHNWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 61
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPDRPWCYVQVGLK 120
DB 62 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPDRPWCYVQVGLK 121
QY 121 PLVQECMVHDCADG 134
DB 122 PLVQECMVHDCADG 135

RESULT 49
US-09-101-272G-96
; Sequence 96, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR

```

FILE REFERENCE: Q50979
CURRENT APPLICATION NUMBER: US/09/101.272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.1
EQ ID NO 96
LENGTH: 201
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: ATFHI-CL chimeric protein
O9-101-272G-96

very Match          34.9%; Score 788; DB 4; Length 201;
'est Local Similarity 100.0%; Pred. No. 2.8e-61;
atches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SNEHLHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKKEGGQHCIDKSKTCYEGNGHFYRG 60
2 SNEHLHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKKEGGQHCIDKSKTCYEGNGHFYRG 61

61 KASDTDTMGRPCLPWN SATVLQOQTYAHRS DALQLGLGKHNYCRNPDNRREPWCYVQVGLX 120
62 KASDTDTMGRPCLPWN SATVLQOQTYAHRS DALQLGLGKHNYCRNPDNRREPWCYVQVGLX 121
121 PLVQECMVHDCADG 134
122 PLVQECMVHDCADG 135

FULT 50
O9-553-498-10
sequence 10, Application US/09553498
tent No. 6309861
GENERAL INFORMATION:
APPLICANT: Ambrosius, Dorothee
APPLICANT: Rudolph, Rainer
APPLICANT: Schaeffner, Joerg
APPLICANT: Schwarz, Elisabeth
TITLE OF INVENTION: Process for the production of naturally folded and secreted prote
FILE REFERENCE: Case 20379
CURRENT APPLICATION NUMBER: US/09/553.498
CURRENT FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: EP99107412.1
PRIOR FILING DATE: 1999-04-26
NUMBER OF SEQ ID NOS: 10
EQ ID NO 10
LENGTH: 378
TYPE: PRT
ORGANISM: E. coli
O9-553-498-10

very Match          34.8%; Score 786; DB 4; Length 378;
'est Local Similarity 44.5%; Pred. No. 9.3e-61;
atches 156; Conservative 47; Mismatches 134; Indels 16; Gaps 6;

50 CYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYAHRS DALQLGLGKHNYCRNPDNR 109
31 CYFGNGSAVRGTHSLTSGASCLPWN SMILLKVITYAQNPESAQLGLGKHNYCRNPDGDA 90
110 RPYCYOVGLKPLVQECMVHDCADGKLPQCG-QKTLRPFKLIIGFEFTIENQNPFAAI 168
91 KPCHWLTNRLTW EYCDVPCS-----TCGLRQYSQPQFRIKGGLFADIASHPWQAAI 144
169 YRHH-RGGSVYVYCGGSLIPCWIGSATHCFIDYPKKEDYIVYLGRSRLNSNTQGMKFE 227
145 FAXHRSRPERFLCGGLILSSCWILSAACHCFQERFPFPHLTVILGRYRVVPEERQKE 204
228 VENLILHKYSAOTLAHNDIALLKIRSEKGRCAQPSRTIQTCLPSWYNDPQFGTSCEI 287
205 VEKYIWHKEFDDDT--YDNDIALQLKSDSRCAQSSVVRTVCLPDLQLOLPDWTCECL 262

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

C: protein - protein search, using sw model
R: On: May 25, 2004, 14:47:10 ; Search time 8.27169 Seconds
(without alignments)
1662.947 Million cell updates/sec

T le: US-09-880-503-8
P: 837
S: 1 SNEHQVPSNCDLNGTGV.....QECMVHDCADGKPSPPPE 143

S: ing table: BLOSUM62
Gapop 10.0 , Gapext 0.5

S: ched: 283366 seqs, 96191526 residues

T: al number of hits satisfying chosen parameters: 283366
M: um DB seq length: 0
M: um DB seq length: 2000000000
P: -processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

D: base : PIR_78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Rank	ID	Score	Query Match %	Description
1	UKHU	837	100.0	u-plasminogen acti
2	UKBAY	757.5	90.5	u-plasminogen acti
3	UKPG	625.5	74.7	u-plasminogen acti
4	JN0560	605	72.3	u-plasminogen acti
5	S18932	592	70.7	u-plasminogen acti
6	UKMS	569	68.0	u-plasminogen acti
7	A34369	335.5	40.1	t-plasminogen acti
8	J30598	334.5	40.1	t-plasminogen acti
9	A35005	334.5	40.0	t-plasminogen acti
10	I38098	328.5	39.2	t-plasminogen acti
11	J30599	328.5	39.2	t-plasminogen acti
12	UKHUT	328.5	38.2	t-plasminogen acti
13	A35029	325	38.8	t-plasminogen acti
14	A39941	315.5	37.7	t-plasminogen acti
15	J30597	310.5	37.1	t-plasminogen acti
16	S28941	268	32.0	coagulation factor
17	J35878	265	31.7	plasma hyaluronan
18	J47495	263.5	31.5	plasma hyaluronan
19	A46688	262	31.3	hepatocyte growth
20	XKHU12	233	27.8	coagulation factor
21	S20600	223	26.6	t-plasminogen acti
22	S45281	221	25.9	coagulation factor
23	A48289	169	20.2	neurotrophic recep
24	B61545	163	19.5	plasmin (EC 3.4.21
25	S00657	162.5	19.4	apoptein(a) (EC
26	A32869	161	19.2	apolipoprotein(a)
27	C61545	160	19.1	plasmin (EC 3.4.21
28	B30848	159.5	19.1	plasmin (EC 3.4.21
29	E61545	159	19.0	plasmin (EC 3.4.21

30	159	19.0	812	1	PLBO	plasmin (EC 3.4.21
31	158.5	18.9	937	2	A45082	neurotrophic recep
32	156.5	18.7	810	1	PLHU	plasmin (EC 3.4.21
33	155	18.5	790	1	PLPG	plasmin (EC 3.4.21
34	155	18.5	812	1	PLMS	plasmin (EC 3.4.21
35	154	18.4	89	2	A60140	plasmin (EC 3.4.21
36	153.5	18.3	810	2	I46260	plasmin (EC 3.4.21
37	152.5	18.2	169	2	A40522	plasmin (EC 3.4.21
38	150	17.9	711	1	A47136	macrophage-stimula
39	149	17.8	455	2	A61545	plasmin (EC 3.4.21
40	148	17.7	943	2	B45082	neurotrophic recep
41	147.5	17.6	716	1	A40332	macrophage-stimula
42	147	17.6	2869	2	T18518	apolipoprotein(a)
43	146.5	17.5	806	2	T18840	hypothetical prote
44	145.5	17.4	716	1	JCS061	macrophage-stimula
45	144.5	17.3	728	1	JH0579	hepatocyte growth
46	142	17.0	411	2	I51285	hepatocyte growth
47	142	17.0	728	1	A35644	hepatocyte growth
48	142	17.0	728	1	A60185	hepatocyte growth
49	141	16.8	710	1	I51283	hepatocyte growth
50	136	16.2	336	2	S33879	plasmin precursor
51	125.5	15.0	618	2	A35827	thrombin (EC 3.4.2
52	124	14.8	622	1	TBHU	thrombin (EC 3.4.2
53	122	14.6	946	1	A47299	ror-related recept
54	117.5	14.0	2352	2	T30201	Notch homolog prot
55	117	14.0	2524	2	A35844	Notch protein - AF
56	115	13.7	825	1	TBRO	thrombin (EC 3.4.2
57	113	13.5	617	2	S10511	hypothetical prote
58	108.5	13.0	1620	2	T27283	notch homolog - se
59	106	12.7	2531	2	T31070	aggreacan precursor
60	105.5	12.6	2109	1	I50421	crumbs protein - f
61	104.5	12.5	2139	2	A35672	Morch B protein -
62	103	12.3	1203	2	A49175	glpi protein precu
63	103	12.3	1295	2	A32901	notch protein homo
64	102	12.2	2531	2	S18188	notch protein homo
65	101.5	12.1	2555	2	A40043	notch protein - fr
66	101.5	12.1	2703	1	A24420	Wnt inhibitory fac
67	101	12.1	379	2	A59180	protein FilC7, 4 [i
68	101	12.1	1722	2	E89753	cell-fate determin
69	101	12.1	2471	2	A49128	MEGFS protein - ra
70	100.5	12.0	1523	2	T13953	Notch A protein -
71	100	11.9	387	2	B49175	Notch-1 protein -
72	100	11.9	2531	2	A46019	Notch homolog Motc
73	99.5	11.9	861	2	A48825	hypothetical prote
74	99.5	11.9	2180	2	T29764	transmembrane prot
75	99.5	11.9	2437	2	S42612	

ALIGNMENTS

RESULT 1
UKHU
u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen ac
N;Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen ac
in form
C;Species: Homo sapiens (man)
C;Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000
C;Accession: A00931; I52209; J70102; A37561; I38102; S65783; A37562; A37563; A37564; A3
R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastiao, G.; Boast, S.; Blasi, F.
Nucleic Acids Res. 13, 2759-2771, 1985
A;Title: The human urokinase-plasminogen activator gene and its promoter.
A;Reference number: A00931; MUID:85215647; PMID:2987867
A;Accession: A00931
A;Molecule type: DNA
A;Residues: 1-431 <RIC>
A;Cross-references: GB:X02419; NID:G37601; PIDN:CAA26268.1; PID:q1834524
A;Note: the authors translated the codon ATG for residue 214 as Ile
R;Nagamine, Y.; Pearson, D.; Grattan, M.
Biochem. Biophys. Res. Commun. 132, 563-569, 1985
A;Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcine
A;Reference number: I52209; MUID:86050639; PMID:3933505

A; accession: I52209
 A; atus: preliminary; translated from GB/EMBL/DBJ
 A; molecule type: DNA
 A; residues: 145-161 <NAG1>
 A; references: GB:K03027; NID:9340174; PIDN:AAA61257.1; PID:9340175
 R; Sai, M.; Hiramatsu, K.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama, G.; 183-188, 1985
 A; title: Molecular cloning of cDNA coding for human preprourokinase.
 A; reference number: J70102; MUID:86056954; PMID:2415429
 A; accession: J70102
 A; molecule type: mRNA
 A; residues: 1-213, 'T', 215-431 <NAG2>
 A; references: GB:K03226; NID:9340155; PIDN:AA97138.1; PID:9340158; GB:D00244; NID:9340158; PID:9340158; PID:9340158
 R; Ge, F.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.
 A; title: Identification and primary sequence of an unspliced human urokinase poly(A) + RN
 A; reference number: A37561; MUID:84272706; PMID:6589620
 A; accession: A37561
 A; molecule type: mRNA
 A; residues: 66-431 <VER>
 A; references: GB:D00244; NID:9220138
 R; Coss, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elsen, D.N.; 139-146, 1985
 A; title: Molecular cloning, sequencing, and expression in *Escherichia coli* of human prep
 A; reference number: I38102; MUID:85203359; PMID:3888571
 A; accession: I38102
 A; atus: preliminary
 A; molecule type: mRNA
 A; residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>
 A; references: EMBL:X02760; NID:935297; PIDN:CAA26535.1; PID:935298
 R; Shimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, B.; 1985
 A; title: Characterization of single chain urokinase-type plasminogen activator with a no
 A; reference number: S65783; MUID:96186279; PMID:8652631
 A; accession: S65783
 A; atus: preliminary
 A; molecule type: mRNA
 A; residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>
 A; references: EMBL:D11143; NID:9311467; PIDN:BAA01919.1; PID:91199928
 R; Zeyler, W.A.; Steffens, G.J.; Otting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L.
 A; title: The primary structure of high molecular mass urokinase from human urine.
 A; reference number: A37562; MUID:83055084; PMID:6754569
 A; accession: A37562
 A; molecule type: protein
 A; residues: 21-177 <GUN>
 R; Haller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O.
 A; title: Human low-molecular-weight urinary urokinase. Partial characterization and prel
 A; reference number: A37563; MUID:83003608; PMID:6749491
 A; accession: A37563
 A; molecule type: protein
 A; residues: 156-176, 179-193, 'T', 195, 'T', 197-224 <SCH>
 R; Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Flohe, L.
 A; title: The complete amino acid sequence of low molecular mass urokinase from human uri
 A; reference number: A37564; MUID:83055039; PMID:6754572
 A; accession: A37564
 A; molecule type: protein
 A; residues: 158-410 <STE>
 R; Atzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K.
 A; title: Carbohydrate composition and presence of a fucose-protein linkage in recombinan
 A; reference number: A35689; MUID:90365737; PMID:2393398
 A; accession: A35689
 A; molecule type: protein
 A; residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>
 A; title: Identification of a fucose and attempt to determine its attachment site
 R; Obani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltz, B.
 A; title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li
 A; reference number: A36697; MUID:91097529; PMID:2125213
 A; accession: A36697

A; molecule type: protein
 A; residues: 21-34 <RAB>
 R; Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.
 A; title: Submitted to the Brookhaven Protein Data Bank, July 1993
 A; reference number: A51255; PDB:KDU
 A; contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R; Li, X.; Smith, R.A.G.; Dobson, C.M.
 A; title: Biochemistry 31, 9562-9571, 1992
 A; reference number: A44375; MUID:93003110; PMID:1327118
 A; contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
 R; Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettlesheim, D.G.; Mazar, A.P.; Olejniczak, Z.
 A; reference number: A66822; PDB:1URK
 A; contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R; Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;
 A; title: Submitted to the Brookhaven Protein Data Bank, July 1995
 A; reference number: A66058; PDB:1LMW
 A; contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175, 179-426
 C; comment: This enzyme is found in urine in a high molecular mass form, consisting of A
 C; comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a
 C; genetics:
 A; gene: GDS:PLAU
 A; cross-references: GDB:119497; OMIM:191840
 A; map position: 10q24-10q24
 A; introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
 C; function:
 A; description: proteolytically activates plasminogen
 A; pathway: fibrinolysis
 C; superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C; keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine proteina
 F; 1-20/Domain: signal sequence #status predicted <SIG>
 F; 21-431/Product: urokinase-type plasminogen activator, single chain form #status predic
 F; 21-177/Product: urokinase-type plasminogen activator chain A #status experimental <MPA
 F; 31-62/Domain: EGF homology <AGF>
 F; 70-151/Domain: kringle homology <KRG>
 F; 156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental <M
 F; 179-431/Product: urokinase-type plasminogen activator chain B #status experimental <MP
 F; 179-419/Domain: trypsin homology <TRY>
 F; 31-39, 33-51, 53-62, 70-151, 91-133, 122-146, 168-299, 209-225, 217-288, 313-382, 345-361, 372-40
 F; 38/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F; 178-179/Cleavage site: lys-ile (plasmin) #status experimental
 F; 224, 275, 376/Active site: His, Asp, Ser #status experimental
 F; 322/Binding site: carbohydrate (Asn) (covalent) #status experimental
 Query Match 100.0%; Score 837; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 9.5e-68;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 SNELHQVPSNCDCLNGTCTVSNKYFNSNHWCNCPKXFGGQHCIDKSKTCYEGNGHFYRG 60
 Db 21 SNELHQVPSNCDCLNGTCTVSNKYFNSNHWCNCPKXFGGQHCIDKSKTCYEGNGHFYRG 80
 Qy 61 KASDTWTGRCLPWNSATVLQOYTHAHRSDALQLGKKNYCNPNRRPWCYVQVGLK 120
 Db 81 KASDTWTGRCLPWNSATVLQOYTHAHRSDALQLGKKNYCNPNRRPWCYVQVGLK 140
 Qy 121 PLVQECNVHDCADGKXPPSPPEE 143
 Db 141 PLVQECNVHDCADGKXPPSPPEE 163
 RESULT 2
 UKBAY
 u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
 C; Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
 C; Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
 C; Accession: S14687; S06651
 R; Au, Y.P.T.; Wang, T.W.; Clowes, A.W.
 Nucleic Acids Res, 18, 3411, 1990
 A; title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminoge
 A; reference number: S14687; MUID:90287734; PMID:2113276
 A; accession: S14687

A Molecule type: mRNA
 A Residues: 1-433 <AUV>
 A Cross-references: EMBL:X51935; NID:938130; PIDN:CAA36200.1.1; PID:938131
 C Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F 1-20/Domain: signal sequence #status predicted <SIG>
 F 1-176/Product: plasminogen activator chain A #status predicted <ACH>
 F 1-61/Domain: EGF homology <EGF>
 F 1-150/Domain: kringle homology <KRG>
 F 1-433/Product: plasminogen activator chain B #status predicted <BCH>
 F 1-421/Domain: trypsin homology <TRY>
 F 17-298,208-224,216-287,315-384,347-363,374-402/Disulfide bonds: #status predicted
 F 13,274,378/Active site: His, Asp, Ser #status predicted
 F 24/Binding site: carbohydrate (Asn) (covalent) #status predicted
 Query Match 90.5%; Score 757.5; DB 1; Length 433;
 Best Local Similarity 92.3%; Pred. No. 1.3e-60; Indels 1; Gaps 1;
 Matches 132; Conservative 4; Mismatches 6;
 Q 1 SNELHGV--PSNCDCINGGTCVSNKYFSNIHWCNCPKFGQHCETDKSKTCYEGNGHFYRG 60
 D 21 SREL-QVPSDCGLNGGTCMSNKYFSNIHWCNCPKFGQHCETDKSKTCYEGNGHFYRG 79
 Q 61 YASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
 D 80 KASTDTMGRSCLAWNSATVLQOTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 139
 Q 121 PLVQECMVHDCADGKPKSPPEE 143
 D 140 QRVCQMVHNCADGKPKSPPEE 162
 R JLT 3
 U 3
 U Plasminogen activator (EC 3.4.21.73) precursor - pig
 N Alternate names: uPA
 C Species: Sus scrofa domestica (domestic pig)
 C Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
 C Accession: A00932
 R Urokinase, Y.; Pearson, D.; Altus, M.S.; Reich, E.
 A ITC: CDNA and gene nucleotide sequence of porcine plasminogen activator.
 A Reference number: A00932; MUID:85087954; PMID:6096832
 A Accession: A00932
 A Molecule type: DNA
 A Residues: 1-240, 'H', 242-442 <NAG1>
 A Experimental source: kidney cell line LIC-PK1
 A Submitted to the Protein Sequence Database, December 1986
 A Reference number: A37566
 A Contents: annotation; correction to residue 241
 C Notes: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
 C Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F 20/Domain: signal sequence #status predicted <SIG>
 F 1-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F 1-64/Domain: EGF homology <EGF>
 F 2-153/Domain: kringle homology <KRG>
 F 90-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F 92/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F 79-310,220-236,228-299,324-393,356-372,383-411/Disulfide bonds: #status predicted
 F 35,286,387/Active site: His, Asp, Ser #status predicted
 Query Match 74.7%; Score 625.5; DB 1; Length 442;
 Best Local Similarity 73.4%; Pred. No. 9.2e-49; Indels 11; Gaps 2;
 Matches 113; Conservative 11; Mismatches 19;
 Q 1 SNELHGV--PSNCDCINGGTCVSNKYFSNIHWCNCPKFGQHCETDKSKTCYEGNGHFY 58
 D 21 SHELHGVESGASNGCGLNGKCVSYKFSNIQRCSCKPKFGQHCETDKSKTCYEGNGHSHY 80

QY 59 RGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVG 118
 DB 81 RKGANTNTGRCPLPWN SATVLNTYHAHRR PDALQLGLGKHNYCRPNDRRRPWCYVQVG 140
 QY 119 LKPLVQECMVHDCADGKPKSPPEE 143
 DB 141 LKQLVQECMVHDCADGKPKSPPEE 174
 RESULT 4
 UN0560
 u-plasminogen activator (EC 3.4.21.73) precursor - bovine
 N Alternate names: uPA
 C Species: Bos primigenius taurus (cattle)
 C Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C Accession: JN0560
 R Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
 Gene 125, 177-183, 1993
 A Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and indu
 A Reference number: JN0560; MUID:93216119; PMID:8385052
 A Accession: JN0560
 A Molecule type: mRNA
 A Residues: 1-433 <KRA>
 A Cross-references: GB:I03546; NID:gl3800; PIDN:AAA51419.1; PID:gl3801
 C Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; tr
 C Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F 1-20/Domain: signal sequence #status predicted <SIG>
 F 21-179/Product: plasminogen activator chain A #status predicted <MAI>
 F 21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F 33-64/Domain: EGF homology <EGF>
 F 72-153/Domain: kringle homology <KRG>
 F 181-433/Product: plasminogen activator chain B #status predicted <MA2>
 F 181-421/Domain: trypsin homology <TRY>
 F 170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
 F 226,277,378/Active site: His, Asp, Ser #status predicted
 Query Match 72.3%; Score 605; DB 1; Length 433;
 Best Local Similarity 72.4%; Pred. No. 6.3e-47;
 Matches 105; Conservative 14; Mismatches 24; Indels 2; Gaps 1;
 QY 1 SNELHGV--PSNCDCINGGTCVSNKYFSNIHWCNCPKFGQHCETDKSKTCYEGNGHFY 58
 DB 21 SNEVHKGESGNSGCLNGKCVTYKFSNIQRCSCKPKFGQHCETDKSKTCYEGNGHSHY 80
 QY 59 RGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVG 118
 DB 81 RKGANDLSGRPCLAWNSATVLLKMTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQIG 140
 QY 119 LKPLVQECMVHDCADGKPKSPPEE 143
 DB 141 LKQFVQFCMVQDCGKSPSPREK 165
 RESULT 5
 SI8932
 u-plasminogen activator (EC 3.4.21.73) precursor - rat
 N Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator
 C Species: Rattus norvegicus (Norway rat)
 C Date: 18-Oct-1989 #sequence_revision 10-Feb-1995 #text_change 18-Jun-1999
 C Accession: S24604; I60186; I53472; SI8932
 R Rabbani, S.A.
 Submitted to the EMBL Data Library, April 1992
 A Reference number: S24604
 A Accession: S24604
 A Molecule type: mRNA
 A Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>
 A Cross-references: EMBL:X65651; NID:g57456; PIDN:CAA46601.1; PID:g57457
 A Experimental source: tissue kidney
 R Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
 Cancer Res. 52, 2489-2496, 1992
 A Title: Transcriptional and posttranscriptional activation of urokinase plasminogen ac
 A Reference number: I60186; MUID:92233409; PMID:1568219
 A Accession: I60186

atus: preliminary; translated from GB/EMBL/DBJ
 A: leucine type: mRNA
 A: sidues: 1-432 <RES>
 A: sss-references: EMBL:X63434; NID:g57465; PIDN:CAA45028.1; PID:g57466
 A: Derimantal source: strain Fischer 344; tissue mammary
 A: mo, P.; Casano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.
 FE: Lett. 306, 193-198, 1992
 A: title: The receptor for the plasminogen activator of urokinase type is up-regulated in
 A: ference number: I53472; MUID:92339549; PMID:1321734
 A: session: I53472
 A: atus: preliminary; translated from GB/EMBL/DBJ
 A: leucine type: DNA
 A: sidues: 31-62 <RE2>
 A: sss-references: EMBL:X66907; NID:g336200; PIDN:CAA47356.1; PID:g938279
 C: netics:
 C: e: ufa
 C: efamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C: words: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F: 19/Domain: signal sequence #status predicted <SIG>
 F: 177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F: -62/Domain: EGF homology <EGF>
 F: 151/Domain: kringle homology <KRG>
 F: 9-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F: 3-420/Domain: trypsin homology <TRY>
 F: 3-300,210-226,218-289,314-383,346-362,373-401/Disulfide bonds: #status predicted
 F: 5,276,377/Active site: His, Asp, Ser #status predicted
 Query Match 70.7%; Score 592; DB 1; Length 432;
 Best Local Similarity 75.6%; Pred. No. 9.2e-46;
 Matches 102; Conservative 10; Mismatches 23; Indels 0; Gaps 0;
 QY 9 SNDCCLNGGTCVSNKYFSNIHWCNCPKFGGHCEDKSKTCYEGNGHFPYRGKASTDTMG 68
 DB 29 SNCCQNGGVCSYKVFSSIRRCSCPKFKGHCEDTSTKTCYHGNGQSYRGKANTDTKG 88
 QY 69 RPLCPWNSATVLOQTYHAHRSALQLGLGKHNYCRPNDRRPPWCYVQGLKPLVQECMV 128
 DB 89 RPLCLMNSPAVLQKPYNAHSPDALSGLGKHNYCRPNDRPQRPWCYVQGLKQFVQECMV 148
 QY 129 HDCADGKKPSSPPEE 143
 DB 149 QDCSLSKKPSSTVDQ 163
 RE AT 6
 UKI plasminogen activator (EC 3.4.21.73) precursor - mouse
 C: eies: Mus musculus (house mouse)
 C: e: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
 C: session: A29420, A24615
 R: jen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
 B: hemistry 26, 8270-8279, 1987
 A: title: The murine urokinase-type plasminogen activator gene.
 A: ference number: A29420; MUID:88163489; PMID:2831940
 A: session: A29420
 A: leucine type: DNA
 A: sidues: 1-433 <DEG>
 A: sss-references: GB:M17922; NID:g202296; PIDN:AAA40539.1; PID:g202297
 R: lin, D.; Vassalli, J.D.; Compeigne, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher,
 Eu J. Blochem. 148, 225-232, 1985
 A: title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase
 A: ference number: A24615; MUID:85179474; PMID:2985383
 A: session: A24615
 A: leucine type: mRNA
 A: sidues: 1-433 <BEL>
 A: sss-references: GB:X02389; NID:g55127; PIDN:CAA26231.1; PID:g55128
 C: netics:
 C: e: uons: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
 C: efamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C: words: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F: 20/Domain: signal sequence #status predicted <SIG>
 F: -178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F: -63/Domain: EGF homology <EGF>

F:71-152/Domain: kringle homology <KRG>
 F:180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F:180-431/Domain: trypsin homology <TRY>
 F:165-301,211-227,218-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
 F:226,277,378/Active site: His, Asp, Ser #status predicted
 Query Match 68.0%; Score 569; DB 1; Length 433;
 Best Local Similarity 70.4%; Pred. No. 1.1e-43;
 Matches 95; Conservative 15; Mismatches 25; Indels 0; Gaps 0;
 QY 9 SNDCCLNGGTCVSNKYFSNIHWCNCPKFGGHCEDKSKTCYEGNGHFPYRGKASTDTMG 68
 DB 30 SNCCQNGGVCSYKVFSSIRRCSCPKFKGHCEDTSTKTCYHGNGDSYRGKANTDTKG 89
 QY 69 RPLCPWNSATVLOQTYHAHRSALQLGLGKHNYCRPNDRRPPWCYVQGLKPLVQECMV 128
 DB 90 RPLCLMNSPAVLQKPYNAHSPDALSGLGKHNYCRPNDRPQRPWCYVQGLKQFVQECMV 149
 QY 129 HDCADGKKPSSPPEE 143
 DB 150 HDCSLSKKPSSTVDQ 164
 RESULT 7
 A34369
 C: Species: Megaderma lyra
 C: plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
 C: Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C: Accession: A34369
 R: Jacob, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob
 J. Biol. Chem. 264, 17947-17952, 1989
 A: Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin
 A: Reference number: A34369; MUID:90036867; PMID:2509450
 A: Accession: A34369
 A: Status: preliminary
 A: Molecule type: mRNA
 A: Residues: 1-477 <GAR>
 A: Cross-references: GB:J05082; NID:g166080; PIDN:AAA31596.1; PID:g166081
 C: Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C: Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F: 1-21/Domain: signal sequence #status predicted <SIG>
 F: 22-36/Domain: propeptide #status predicted <PRO>
 F: 37-477/Product: plasminogen activator #status predicted <PLA>
 F: 42-79/Domain: fibronectin type I repeat homology <1FA>
 F: 128-209/Domain: kringle homology <KRG>
 F: 226-471/Domain: trypsin homology <TRY>
 F: 42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
 F: 272,321,428/Active site: His, Asp, Ser #status predicted
 Query Match 40.1%; Score 335.5; DB 1; Length 477;
 Best Local Similarity 46.3%; Pred. No. 1e-22;
 Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;
 QY 3 ELHOVP---SNCDCLNGGTCVSNKYFSNIHWCNCPKFGGHCEDKSKTCYEGNGHGFY 58
 DB 78 QCHTVPKVCSSELRFCFNGHCWQAASFSDP-VQCQPKGYTGKQCEVDTHATCKDQGVY 136
 QY 59 RGKASTDTMGRCPLMNSATVLOQTYHAHRSALQLGLGKHNYCRPNDRRPPWCYVQGL 118
 DB 137 RGTWSTSESQAQCIWNNSNLLTERTYNGRSDAITLGLGNHNYCRPNDRPNNKPKWCYVKA 196
 QY 119 LKPLVQECMVHDC 132
 DB 197 SKFILEFCSPVPCS 210
 RESULT 8
 JS0598
 t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
 N: Alternate names: tissue plasminogen activator
 C: Species: Desmodus rotundus (common vampire bat)
 C: Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999

C; words: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F; 21/Domain: signal sequence #status predicted <SIG>
F; 36/Domain: propeptide #status predicted <PRO>
F; 431/Product: plasminogen activator beta #status predicted <PLA>
F; 74/Domain: EGF homology <EGF>
F; 163/Domain: kringle homology <KR>
F; 425/Domain: trypsin homology <TRY>
F; 52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide bond
F; 3352/Binding site: carbohydrate (Aen) (covalent) #status predicted
F; 9-180/Cleavage site: His-Ser (plasmin) #status predicted
F; 5,275,382/Active site: His, Asp, Ser #status predicted
F; 5-361,378-406/Disulfide bonds: #status predicted
F; 39.2%; Score 328.5; DB 2; Length 431;
F; Local Similarity 47.6%; Pred. No. 4.1e-22;
F; Conservative 16; Mismatches 48; Indels 1; Gaps 1;
QY 9 SNDCCLNGTCSNKFYSIHWCNCPKFGGQHCIDSKTCYEGNGHYRKGASTDTMG 68
Db 42 SELRCFNGGTQCAASFDF-VCQPKGYTGQCEVDTHATCYKQGVYRGTSSTSG 100
QY 69 RPLCPNNSATVLOQTHARSALQGLGKHYCRNPNRRPWCYVQVGLKPLVQECW 128
Db 101 AQCIWNNSLLTRTYNGRSDAITLGLGNHYCRNPNNSRWCYVVKASKFILEFCV 160
QY 129 HDCA 132
Db 161 PVCS 164
RE 12
UK 1
t- aminogen activator (EC 3.4.21.68) precursor (validated) - human
N; Bernate names: t-PA; tissue plasminogen activator
C; species: Homo sapiens (man)
F; 14-Nov-1983 #sequence revision 14-Nov-1983 #text_change 08-Dec-2000
C; Session: A94004; A23529; JT0562; A93293; S02125; A91343; A93951; A91322; A54645; 160
R; T.; Elgh, P.; Lund, B.
F; Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A; Title: The structure of the human tissue-type plasminogen activator gene: correlation
A; Reference number: A94004; MUID:84298137; PMID:6089198
A; Session: A94004
A; Molecule type: DNA
A; Residues: 1-562 <NT>
A; Cross-references: GB:L00141
A; Note: The codon given for residue 93 (ACC) is inconsistent with the authors' translation
R; Lezner Degen, S.J.; Rajput, B.; Reich, E.
J. Biol. Chem. 261, 6972-6985, 1986
A; Title: The human tissue plasminogen activator gene.
A; Reference number: A23529; MUID:86196143; PMID:3009482
A; Session: A23529
A; Molecule type: DNA
A; Residues: 1-562 <DEG>
A; Cross-references: GB:K03021; NID:G339817; PIDN:AAA98809.1; PID:G339818
R; Igaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.
A; Biol. Chem. 55, 1225-1232, 1991
A; Title: Purification and characterization of tissue plasminogen activator secreted by H
A; Session: JT0562; MUID:91291340; PMID:1368681
A; Molecule type: mRNA
A; Residues: 31-562 <ITA>
A; Cross-references: DBJ:001096; NID:G220128; PIDN:AAA00881.1; PID:G441174
A; Experimental source: embryonic lung fibroblast IMR-90 cells
A; Note: part of this sequence, including the amino end of the mature protein, was confir
R; Ulica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vohar, G.A.; Ward, C.A.; Bennett
Na 301, 214-221, 1983
A; Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
A; Reference number: A93293; MUID:83115262; PMID:6337343
A; Session: A93293
A; Molecule type: mRNA
A; Residues: 1-562 <PEN>
A; Cross-references: GB:L00141
A; Experimental source: melanoma cells

R;Sasaki, H.; Saito, Y.; Hayaashi, M.; Otsuka, K.; Niwa, M.
Nucleic Acids Res. 16, 5695, 1988
A; Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
A; Reference number: S02125; MUID:88262579; PMID:3133640
A; Accession: S02125
A; Status: translation not shown
A; Molecule type: mRNA
A; Residues: 1-562 <SAS>
A; Cross-references: EMBL:X07393; NID:G37243; PIDN:CAA30302.1; PID:G37244
A; Experimental source: fetal lung cells
R;Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma
FEBS Lett. 189, 145-149, 1985
A; Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
A; Reference number: A91343; MUID:85285620; PMID:3896853
A; Accession: A91343
A; Molecule type: mRNA
A; Residues: 1-38, 'G', '86-433, 'E', 435-562 <KAG>
A; Experimental source: Detroit 562 cells; ATCC 138
R;Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A; Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activ
A; Reference number: A93951; MUID:83169656; PMID:6572897
A; Accession: A93951
A; Molecule type: mRNA
A; Residues: 251-358 <EDL>
A; Experimental source: melanoma cells
R;Pohl, G.; Kaliström, M.; Bergsdorf, N.; Wallen, P.; Jörnvall, H.
Biochemistry 23, 3701-3707, 1984
A; Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
differences.
A; Reference number: A90488; MUID:85000468; PMID:6433976
R;Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jörnvall, H.
FEBS Lett. 168, 29-32, 1984
A; Title: Differences between uterine and melanoma forms of tissue plasminogen activator.
A; Reference number: A91322; MUID:84158956; PMID:6538514
A; Accession: A91322
A; Molecule type: protein
A; Residues: 33-45;311-320 <POH>
A; Experimental source: uterus
A; Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
R;Biol. Chem. 261, 14214-14218, 1986
A; Reference number: A37567; MUID:87033611; PMID:3021732
R;Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Enger
EMBO J. 5, 3525-3530, 1986
A; Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen ac
A; Reference number: A37568; MUID:87161761; PMID:3030730
A; Contents: annotation; fibrin binding site
R;Dodd, I.; Nunn, B.; Robinson, J.H.
Thromb. Haemost. 59, 523-528, 1988
A; Title: Isolation, identification and pharmacokinetic properties of human tissue-type p
A; Reference number: A60902; MUID:89044681; PMID:3142086
A; Contents: annotation; novel forms of expressed recombinant t-PA
R;Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Entage, J.S.; Opdenakker, G.;
Mol. Biol. Med. 3, 279-292, 1986
A; Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expr
A; Reference number: A54645; MUID:86284200; PMID:3090401
A; Accession: A54645
A; Molecule type: mRNA
A; Residues: 1-562 <HAR>
A; Cross-references: GB:M15518; NID:G190031; PIDN:AAA60111.1; PID:G190032
A; Note: parts of this sequence were confirmed by peptide sequencing
R;Reddy, V.B.; Garramone, A.J.; Sasak, H.; Wei, C.
DNA 6, 461-472, 1987
A; Title: Expression of human uterine tissue-type plasminogen activator in mouse cells us
A; Reference number: I60110; MUID:88054470; PMID:2824147
A; Accession: I60110
A; Status: translated from GB/EMBL/DBJ
A; Molecule type: mRNA
A; Residues: 1-562 <RES>
A; Cross-references: GB:M18182; NID:G340176; PIDN:AAA36800.1; PID:G340177

Risher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
 J Biol. Chem. 260, 11223-11230, 1985
 A title: Isolation and characterization of the human tissue-type plasminogen activator cDNA
 A reference number: 155233; PMID:85289338; PMID:3161893
 A accession: 155232
 A status: preliminary; translated from GB/EMBL/DBJ
 A molecule type: DNA
 A residues: 1-36 <R2>
 A cross-references: GB:M11890; NID:G339837; PIDN:AAA61213.1; PID:G339839
 C comment: Cleavage by plasmin or trypsin produces two chains held together by a single
 C comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond. It
 C comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat.
 C enetics:
 A gene: GDB:PLAT
 A cross-references: GDB:119496; OMIM:173370
 A map position: 8p12-8p12
 A atoms: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 510
 C superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
 F 4-23/Domain: signal sequence #status predicted <SIG>
 F 4-32/Domain: propeptide #status predicted <PRO>
 F 3-562/Product: t-plasminogen activator #status experimental <MAT>
 F 1-78/Domain: fibronectin type I repeat homology <I>
 F 5-119/Domain: EGF homology <EGF>
 F 27-208/Domain: kringle homology <KR1>
 F 15-296/Domain: kringle homology <KR2>
 F 11-556/Domain: t-plasminogen activator chain B #status experimental <BCH>
 F 11-562/Product: t-plasminogen activator chain B #status experimental <BCH>
 F 1-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203,215-296,236-278,267-291,299-4
 F 2,483/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F 19/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
 F 10-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental
 F 37,406/Active site: His, Asp #status predicted
 F 13/Active site: Ser #status experimental
 Query Match 39.2%; Score 328.5; DB 1; Length 562;
 Best Local Similarity 46.3%; Pred. No. 5.1e-22;
 Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;
 Q 3 ELHQP-SNCD---CLNGTGVSNKVFNSIHWCNCPKFGQHCIDKSKTCYEGNGHFY 58
 D 77 QCHSVFVRSCEPRFCNGTCQQAALYFSDP-VCCQPEGFAGKCCIDTRATCFEGQGIT 135
 Q 59 RGKASDTMGRCPLPNSATVLQOTVHAHRSALQLGLGKHNYCRPNRRPWCYVQVG 118
 D 136 RGTWSTAEGASCTWNSSALAQKYSGRPRPRLGLGNHNYCRPNRRPDRSKPCYVFK 195
 Q 119 LKPLVQECMVHCADG 134
 D 196 GYSSEFCSTPACSEG 211
 R ULT 13
 A 129
 C plasminogen activator (EC 3.4.21.68) precursor - rat
 C species: Rattus norvegicus (Norway rat)
 C date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C accession: A35029; A31597
 R ang, P.; Ohlsson, M.; Ny, T.
 J Biol. Chem. 265, 2022-2027, 1990
 A title: The structure of the TARA-less rat tissue-type plasminogen activator gene. Spec
 A reference number: A35029; PMID:190130448; PMID:2105315
 A accession: A35029
 A status: preliminary
 A molecule type: DNA
 A residues: 1-559 <PEN>
 A cross-references: GB:M1197; NID:G207429; PIDN:AAA42261.1; PID:G207431; GB:J05226
 F 7, 671-677, 1986
 F 7; T.; Leonardsson, G.; Huseh, A.J.W.
 D title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator
 A reference number: A31597; PMID:89170114; PMID:3149445
 A accession: A31597

A: Molecule type: mRNA
 A: Residues: 1-379; 'K', 381-559 <NVT>
 A: Cross-references: GB:M23697; NID:G530159; PIDN:AAA1812.1; PID:G530160
 C: Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho
 C: Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F 1-17/Domain: signal sequence #status predicted <SIG>
 F 18-29/Domain: propeptide #status predicted <PRO>
 F 30-559/Product: t-plasminogen activator #status predicted <MAT>
 F 30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
 F 38-75/Domain: fibronectin type I repeat homology <I>
 F 83-116/Domain: EGF homology <EGF>
 F 124-205/Domain: kringle homology <KR1>
 F 213-294/Domain: kringle homology <KR2>
 F 309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
 F 309-553/Domain: trypsin homology <TRY>
 F 38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-
 F 149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F 308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
 F 355,404,510/Active site: His, Asp, Ser #status predicted
 Query Match 38.8%; Score 325; DB 1; Length 559;
 Best Local Similarity 44.1%; Pred. No. 1e-21;
 Matches 64; Conservative 15; Mismatches 56; Indels 10; Gaps 3;
 Q 3 ELHQP-----SNCDCLNGTGVSNKVFNSIHWCNCPKFGQHCIDKSKTCYEGNGHFY 58
 D 74 QCHSVFVRSCEPRFCNGTCQQAALYFSDP-VCCQPEGFAGKCCIDTRATCFEGQGIT 132
 Q 59 RGKASDTMGRCPLPNSATVLQOTVHAHRSALQLGLGKHNYCRPNRRPWCYVQVG 118
 D 133 RGTWSTAEGASCTWNSSALSKPYSAKPPNAIKLGLGNHNYCRPNRRPDRSKPCYVFK 192
 Q 119 LKPLVQECMVHCADGCKPSPPEE 143
 D 193 GYKTEFCSTPAC-----PKGPTED 212
 RESULT 14
 A29941
 t-plasminogen activator (EC 3.4.21.68) precursor - mouse
 C: Species: Mus musculus (house mouse)
 C: Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C: Accession: A29941; S48205; S48207; S48206
 R: Rickles, R.J.; Darrow, A.L.; Strickland, S.
 J. Biol. Chem. 263, 1563-1569, 1988
 A: Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator m
 A: Reference number: A29941; PMID:88087303; PMID:2826484
 A: Accession: A29941
 A: Molecule type: mRNA
 A: Residues: 1-559 <RIC>
 R: Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
 Eur. J. Biochem. 224, 863-871, 1994
 A: Title: Characterization of the murine plasma fibrinolytic system.
 A: Reference number: S48202; PMID:95010076; PMID:7523120
 A: Accession: S48205
 A: Molecule type: protein
 A: Residues: 33-37, 'X', 39-40 <LIJ>
 A: Accession: S48207
 A: Molecule type: protein
 A: Residues: 309-316 <LIJ>
 A: Accession: S48206
 A: Molecule type: protein
 A: Residues: 33-37, 'X', 39-40 <LIW>
 C: Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho
 C: Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F 1-17/Domain: signal sequence #status predicted <SIG>
 F 18-29/Domain: propeptide #status predicted <PRO>
 F 30-559/Product: t-plasminogen activator #status predicted <MAT>
 F 30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
 F 38-75/Domain: fibronectin type I repeat homology <I>
 F 83-116/Domain: EGF homology <EGF>
 F 124-205/Domain: kringle homology <KR1>

coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)
N/Alternate names: Hageman factor
C/Species: Cavia porcellus (guinea pig)
C/Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 21-Jan-2000
C/Accession: S28941
R/Semba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.; Okabe, H.
Biochim. Biophys. Acta 1159, 113-121, 1992
A/Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage site
A/Reference number: S28941; MUID:93003367; PMID:1390917
A/Accession: S28941
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-603 <SEM>
A/Cross-references: EMBL:X68615; NID:G49578; PIDN:CAA48600.1; PID:G49579
C/Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C/Keywords: hydrolase; serine proteinase
F/45-87/Domain: fibronectin type II repeat homology <FBI>
F/134-169/Domain: fibronectin type I repeat homology <FBI>
F/177-208/Domain: EGF homology <EGF>
F/216-294/Domain: kringle homology <KRG>
F/359-597/Domain: trypsin homology <TRY>

Query Match 32.0%; Score 268; DB 2; Length 603;
Best Local Similarity 38.6%; Pred. No. 1.4e-16;
Matches 51; Conservative 21; Mismatches 52; Indels 8; Gaps 3;

Qy 13 CLINGGTCVKNKVFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRKGASTDTWGRPCL 72
Db 182 CLNGGRCLR---VEGHLLCDECPMGVTPFCDDLDTTASCYEGRGVSYRGMAATTTVSGAKCQ 238
Qy 73 PWSNATVLOQTYAHNRSD-ALQLGLGKXNYCRNPNNRRPWCYVQGLKPLVQECMVHDC 131
Db 239 RWAS-----EATYNNMTAEQALRGRLGHTTCFNPNDNDRPWC FVMWGNLSWEYCDLAQC 294
Qy 132 ADGKKPSSPPEE 143
Db 295 QYPPQPTATPHD 306

RESULT 17
JC5878
plasma hyaluronan-binding protein precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999
R/Accession: JC5878
R/Hashimoto, K.; Tobe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, N.
Biol. Pharm. Bull. 20, 1127-1130, 1997
A/Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-bin
A/Reference number: JC5878; MUID:98065239; PMID:9401717
A/Accession: JC5878
A/Molecule type: mRNA
A/Residues: 1-558 <HAS>
C/Comment: This protein acts as serine protease.
C/Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsi
F/1-23/Domain: signal sequence #status predicted <SIG>
F/24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <MATL>
F/75-106/Domain: EGF homology <EG1>
F/113-145/Domain: EGF homology <EG2>
F/152-185/Domain: EGF homology <EG3>
F/152-274/Domain: kringle homology <KRI>
F/312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MATS
F/312-548/Domain: trypsin homology <TRY>

Query Match 31.7%; Score 265; DB 2; Length 558;
Best Local Similarity 41.1%; Pred. No. 2.5e-16;
Matches 53; Conservative 19; Mismatches 51; Indels 6; Gaps 4;

Qy 13 CLINGGTCVKNKVFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRKGASTDTWGRPCL 72
Db 157 CQNGGVCSSHRRSRF-TACCPDQYKGFCEITGPD-DCYVGDGYSYRGVSKTVNQNFCL 214
Qy 73 PWSNATVLOQTYAHNRSDALQLGLGKXNYCRNPNNRRPWCYVQGLKPLVQECMVHDC 131

D 215 YWNSHLLQETYNMFEDAEHTGIAEHNFRCRPNQDGHKWCFCVKNSEKVKWEYCDVTVC 274
Q 132 ADGKKPSSP 140
D 275 ---PVPDFP 280
R JLT 18
P sma hyaluronan-binding protein precursor - human
N lternate names: hepatocyte growth factor activator-like protein; PHBP
N tains: serine proteinase (SC 3.4.21.-)
C species: Homo sapiens (man)
C cession: J04795
E 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 19-Jul-2002
F 101-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, T.; Tomita, M.
J Biochem. 119, 1157-1165, 1996
A title: Purification and characterization of a novel hyaluronan-binding protein (PHBP)
I activator.
A reference number: J04795; MUID:96425001; PMID:8827452
A cession: J04795
A cecule type: mRNA
A residues: 1-560<CHO>
A cross-references: GB:S83182; NID:g1836158; PIDN:AAB46909.1; PID:g1836159
A perimental source: plasma
A te: parts of this sequence, including the amino ends of the mature chains, were dete
C netics:
A ne: GDB:HABP2; HABP; PHBP; HGFPAL
A cross-references: GDB:4573962
C mplex: a disulfide-bonded heterodimer of chains produced from the same precursor; th
C perfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypt
C ywords: chondroitin sulfate proteoglycan; glycoprotein; hyaluronic acid; hydrolase;
F -23/Domain: signal sequence #status predicted <SIG>
F 4-113/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>
F 7-108/Domain: EGF homology <EG1>
F 15-147/Domain: EGF homology <EG2>
F 34-187/Domain: EGF homology <EG3>
F 34-276/Domain: kringle homology <KRI>
F 14-550/Domain: tryptsin homology <TRY>
F 14-516/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted
F 1-207/Binding site: carbohydrate (Asn) (covalent) #status predicted
F 1-88,82-97,99-108,115-125,120-136,138-147,154-165,159-176,178-187,194-276,215-257,248
F 2-405,509/Active site: His, Asp, Ser #status predicted
Query Match 31.5%; Score 263.5; DB 1; Length 560;
Best Local Similarity 41.0%; Pred. No. 3.4e-16;
Matches 55; Conservative 18; Mismatches 56; Indels 5; Gaps 4;
Q 13 CLNGGTCVSNKYFNIHWCNCPKFGGQHCIDSKYCYEGNGHFGKASTDTMGRPCL 72
D 159 CQNGATCSRHKRRSKF-TCACPDFGKFCBEG-SDDCYVGDGYSYRGKMRVTYNQACL 216
Q 73 PWSATVLSQTYHAHRSDALQLGLGKHCNCRPNRRRPWCYOVGLKPLVQE-CMVHDC 131
D 217 YWNSHLLQETYNMFEDAEHTGIAEHNFRCRPNQDGHKWCFCVKNSEKVKWEYCDVSAC 276
Q 132 A--DCKFPSSPPEE 143
D 277 SAQDVAYPEESFTE 290
R JLT 19
A loocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human
C species: Homo sapiens (man)
C ate: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
C cession: A46688
F yazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
J Biol. Chem. 268, 10024-10028, 1993
A title: Molecular cloning and sequence analysis of the cDNA for a human serine protease
C gulation factor XII.
A ference number: A46688; MUID:93252878; PMID:7683665
A cession: A46688
A;Molecule type: mRNA
A;Residues: 1-655 <MIY>
A;Cross-references: DDBJ:D14012; NID:g219680; PIDN:BAA03113.1; PID:g219681
A;Experimental source: liver (mRNA); serum (protein)
A;Note: sequence extracted from NCBI backbone (NCBIN:131227, NCBI:P:131228)
A;Note: parts of the sequence, including the amino ends of the heavy and light chains,
C;Genetics:
A;Gene: GDB:HGFA; HGFA; HGFAF
A;Cross-references: GDB:9954514
A;Map position: 4p16-4p16
A;Function:
A;Description: activates hepatocyte growth factor by specific proteolytic cleavage
A;Pathway: tissue repair and regeneration
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology
C;Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
F;1-34/Domain: signal sequence #status predicted <SIG>
F;108-148/Domain: fibronectin type II repeat homology <IF2>
F;164-197/Domain: EGF homology <EG1>
F;202-237/Domain: fibronectin type I repeat homology <IF1>
F;245-278/Domain: EGF homology <EG2>
F;286-367/Domain: kringle homology <KRG>
F;373-407/Product: hepatocyte growth factor activator light chain #status experimental
F;408-655/Product: hepatocyte growth factor activator heavy chain #status experimental
F;40-48,290,468,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;164-175,168-186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-3
F;447,497,598/Active site: His, Asp, Ser #status predicted
Query Match 31.3%; Score 262; DB 1; Length 655;
Best Local Similarity 36.9%; Pred. No. 5.4e-16;
Matches 58; Conservative 12; Mismatches 57; Indels 30; Gaps 3;
QY 5 HQVPSNCDCLNGGTCVSNKYFNIHWCNCPKFGGQHCIDSKYCYEGNGH 56
DB 242 HTACUSSPCLNGGTC-----HLIVATGTTVCACPPGFAGRLCNIEPDERCFIGNGT 292
QY 57 FYRGKASTDTMGRPCLPMNSATVLSQTYHAHRSDALQLGLGKHCNCRPNRRRPWCYQV 116
DB 293 GYRGVASTSASGLSCLANNSDLLYQLHVDVSGAALGLGPHACRPNDRPWCYV 352
QY 117 VGLKPLVQECMVHDC-----ADGKKPSSP 140
DB 353 KDSALSWEYCRLEACESLTVQLSPDLLATLPEPASP 389
RESULT 20
KFUHL2
Coagulation factor XIIa (EC 3.4.21.38) precursor [validated] - human
N;Alternate names: Hageman factor (activated)
C;Species: Homo sapiens (man)
C;Date: 27-Nov-1985 #sequence_revision 30-Jun-1991 #text_change 08-Dec-2000
C;Accession: A29411; A26814; A00930; A25191; A22248; A21037
R;Cool. D.E.: MacGillivray, R.T.A.
J. Biol. Chem. 262, 13662-13673, 1987
A;Title: Characterization of the human blood coagulation factor XII gene. Intron/exon 9
A;Reference number: A29411; MUID:88007593; PMID:2888762
A;Accession: A29411
A;Molecule type: DNA
A;Residues: 1-615 <COO>
A;Cross-references: GB:M17466; GB:J02807; NID:g180355; PIDN:AAB59490.1; PID:g180357
R;Tripos, M.; Citarella, F.; Guida, S.; Galeffi, P.; Fantoni, A.; Cortese, R.
Nucleic Acids Res. 14, 3146, 1986
A;Title: cDNA sequence coding for human coagulation factor XII (Hageman).
A;Reference number: A26814; MUID:86176794; PMID:3754331
A;Accession: A26814
A;Molecule type: mRNA
A;Residues: 4-615 <TRI>
A;Cross-references: GB:M31315; NID:g182291; PIDN:AAA70225.1; PID:g182292
R;Cool, D.E.; Eggell, C.J.S.; Louie, G.V.; Zoller, M.J.; Brayer, G.D.; MacGillivray, R.
J. Biol. Chem. 260, 13666-13676, 1985
A;Title: Characterization of human blood coagulation factor XII cDNA. Prediction of the
A;Reference number: A00930; MUID:86033830; PMID:3877053
A;Accession: A00930

JS0600

t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0600
R:Kraatzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0600
A:Molecule type: mRNA
A:Residues: 1-394 <KRA>
A:Cross-references: GB:M63990; MID:g166078; PIDN:AAA31595.1; PID:g166079
A:Note: the authors translated the codon ATC for residue 75 as Thr
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-394/Product: plasminogen activator gamma #status predicted <PLA>
F:45-126/Domain: kringle homology <KRG>
F:143-388/Domain: trypsin homology <TRY>
F:45-126,66-108,97-121,131-282,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:
F:142-143/Cleavage site: His-Ser (plasmin) #status predicted
F:189,238,345/Active site: His, Asp, Ser #status predicted
F:315/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 26.6%; Score 223; DB 2; Length 394;
Best Local Similarity 44.3%; Pred. No. 1.1e-12;
Matches 39; Conservative 13; Mismatches 36; Indels 0; Gaps 0;

QY 45 DKSCTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRN 104
DB 40 DPHATCYKQGVYRTGTWSTSGGAQCINMNSNLLIRTYNGRMPEAVKLGNGHNYCRN 99

QY 105 PDNRPRPCYVQVGLKPLVQECMWHDC A 132
DB 100 PDGASKPCYVIXARKFTSCSVFVCS 127

RESULT 22

S45281

coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)
N:Alternate names: Hageman factor (activated)
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C:Accession: S45281; A61329
R:Shibuya, Y.; Semba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A:Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): comp
A:Reference number: S45281; MUID:94242782; PMID:8186251
A:Accession: S45281
A:Molecule type: mRNA
A:Residues: 1-593 <SHI>
A:Cross-references: GB:S70164
A:Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70 as
s Pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 as Gl
is, and ATC for residue 505 as Leu
R:Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A:Title: Isolation and characterization of bovine factor XII (Hageman factor).
A:Reference number: A61329; MUID:77182112; PMID:661210
A:Accession: A61329
A:Molecule type: protein
A:Residues: 10-16,'X',18-19,525-550 <FUJ>
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma;
F:37-78/Domain: fibronectin type II repeat homology <IF2>
F:88-120/Domain: EGF homology <EGF>
F:125-150/Domain: fibronectin type I repeat homology <FBI>
F:207-287/Domain: kringle homology <KRG>
F:350-587/Domain: trypsin homology <TRY>
F:541/Active site: Ser #status predicted

```
F:350-38//Domain: cyprusin homology <iki>
F:541/Active site: Ser #status predicted
```



```

very Match      25.9%; Score 217; DB 2; Length 593;
st Local Similarity 35.7%; Pred. No. 5.4e-12;
atches 46; Conservative 15; Mismatches 58; Indels 10; Gaps 4;

Q   6 QVPENCCLAGGTCVSNKYFNIIHWNCPPKFGGCHCEIDKSCTCYE--GNGHFYRGKAS 63
D   166 QVCNTNPLNGDSCLQAE---GHLRCAPSFAGRLCDVLKASCYDDRDRLGLSYRGMAG 222
Q   64 TDTMGRPCLPWNSATVLQQTY-HAHRSDALQLGLGKINYCRNPDRRRPWCYYQVGILKPL 122
D   223 TTLSGAPCQSASWAS---EATVYNVTASOVNLNWGLGDHAFCRNPDPNDTRPWCFTWKGDRLS 278
Q   123 VQECCWHDC 131
D   279 WNYCRLAPC 287

R   ILT 23
A   289
N   trophic receptor ror precursor - fruit fly (Drosophila melanogaster)
N   lternate names: trk-related receptor
N   ntains: protein-tyrosine kinase (EC 2.7.1.112)
N   ecies: Drosophila melanogaster
N   te: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 17-Nov-2000
N   ccession: A48289
R   lson, C.; Goezrdhan, D.C.I.; Steller, H.
F   . Natl. Acad. Sci. U.S.A. 90, 7103-7113, 1993
A   ticle: Dror, a potential neurotrophic receptor gene, encodes a Drosophila homolog of tyro
A   eference number: A48289; MUID:93348222; PMID:8394009
A   ccession: A48289
A   lecule type: mRNA
A   aidues: 1-685 <WU>
A   oss-references: GB:L20297; NID:g348103; PIDN:AAA28860.1; PID:g348104
C   netics:
A   ne: FlyBase:bek
A   oss-references: FlyBase:FBgn0010407
C   perfamily: Glycophila neurotrophic receptor ror; kringle homology; protein kinase ho
C   ywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyro
F   17-310/Domain: kringle homology <KRG>
F   14-338/Domain: transmembrane #status predicted <TML>
F   18-677/Domain: protein kinase homology <KIN>
F   16-424/Region: protein kinase ATP-binding motif
F   1,63,129,144,250/Binding site: carbohydrate (Asn) (covalent) #status predicted

very Match      20.2%; Score 169; DB 1; Length 685;
st Local Similarity 28.8%; Pred. No. 1.2e-07;
atches 38; Conservative 27; Mismatches 35; Indels 32; Gaps 6;

Q   3 ELHVPSNCDCLANGTCVSNNKYFSNIHWNCPPKFGGCHCEIDKSCTCYEGNHFFYRGKA 62
D   212 DCQKLPHQKDCLSLIGITI-----EVDKTENCYWEDGSTYRGVA 249
Q   63 STDYMGPRCLPWNSATVLQQTYHAHRSDALQLGLGKINYCRNPDRRRPWCYYQVGILKP 121
D   250 NVSASGRPCLRW---SWLMKEI-----SDFPEL-IGQ-NYCRNPGSVENSPWCFVSSSRER 300
Q   122 LVQECMVHDCAD 133
D   301 IIELCDIPKCAD 312

R   ILT 24
B   345
P   min (EC 3.4.21.7) precursor - sheep (fragments)
N   lternate names: plasminogen
N   ontains: miniplasminogen
C   ecies: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C   te: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 17-Mar-1999
C   ccession: B61545; S28200
R   hallier, J.; Rickii, E.E.
E   me 40, 63-69, 1988

```

A>Title: Structural aspects of the plasminogen of various species.
A|Reference number: A6154S; MUID:8900501S; PMID:316897S
A|Accession: B6154S
A|Molecule type: protein
A|Residues: 1-37;38-117 <SCH>
R|Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. S. 21-25, 1992
A|Title: Complete amino acid sequence of ovine miniplasminogen.
A|Reference number: S28200; MUID:9314999S; PMID:1492092
A|Accession: S28200
A|Molecule type: protein
A|Residues: 118-450 <SC2>
C|Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolog
C|Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase;
F:1-37;38-117,118-460/Product: plasminogen (fragments) #status experimental <PRO>
F:1-37/Domin: activation peptide (fragment) #status experimental <APT>
F:38-117,118-230,231-460/Product: plasmin (fragments) #status experimental <MAT>
F:41-118/Domin: kringle homology <KR4>
F:118-460/Product: miniplasminogen #status experimental <MIN>
F:132-211/Domin: kringle homology <KR5>
F:226-460/Domin: plasmin chain B #status experimental <BCB>
F:231-453/Domin: trypsin homology <TRY>
F:272,315,410/Active site: His, Asp, Ser #status predicted

Query Match 19.5%; Score 163; DB 2; Length 450;
Best Local Similarity 33.6%; Pred No. 3e-07; Gaps 6;
Matches 40; Conservative 16; Mismatches 43; Indels 20;

QY 35 KFFGGQ-----HCEIDKSKTCYEGNHFGYRGKASTDMGRPCLPWNSATV--LQQTYY 85
Db 20 KLAKRSVEDCAAKE-BEADQCYHGNGQGYRGTSSTVTGKCKQSWSMIPHRHQKTPE 78
QY 86 AHRSDALQLGLGHNYCNPNRRPNCVCQVGLKLPLV--QECMVHDCA DGKKSPPE 142
Db 79 STPNAGLTW----NYCRNPDADRKPWCYT---TPRVWFECNUKKAFAQPSVENPPPE 129

RESULT 25
S00657
apoprotein(a) (EC 3.4.21.-) precursor [validated] - human
N|Alternate names: apolipoprotein(a); lipoprotein(a) chain apo(a)
C|Species: Homo sapiens (man)
C|Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text change 08-Dec-2000
C|Accession: S00657; A28017; A47277; I60906; A47233; I52415; I65286
R:McLean, J.W.; Tomlinson, J.E.; Kuang, W.J.; Eaton, D.L.; Chen, E.Y.; Fless, G.M.; Sca
Nature 330, 132-137, 1987
A|Title: cDNA sequence of human apolipoprotein(a) is homologous to plasminogen.
A|Reference number: S00657; MUID:88039109; PMID:3670400
A|Accession: S00657
A|Molecule type: mRNA
A|Residues: 1-4548 <MC1>
A|Cross-references: GB:X06290; EMBL:X06696; NID:g28619; PIDN:CAA29618.1; PID:g28620
R:Eaton, D.L.; Fless, G.M.; Kohr, W.J.; McLean, J.W.; Xu, Q.T.; Miller, C.G.; Lawn, R.M.
Proc Natl. Acad. Sci. U.S.A. 84, 3224-3228, 1987
A|Title: Partial amino acid sequence of apolipoprotein(a) shows that it is homologous t
A|Reference number: A28017; MUID:87204109; PMID:3472206
A|Accession: A28017
A|Molecule type: protein
A|Residues: 20-21,'P',23-34;177-179,'N',181-186,'T',188-196,'DKG',200;292-314,'W',316-3
X' 4396-4401 <EAT>
R,Wade, D.P.; Clarke, J.G.; Lindahl, G.E.; Liu, A.C.; Zysow, B.R.; Meer, K.; Schwartz,
Proc. Natl. Acad. Sci. U.S.A. 90, 1369-1373, 1993
A|Title: 5' control regions of the apolipoprotein(a) gene and members of the related pl
A|Reference number: A47277; MUID:93165698; PMID:7679504
A|Accession: A47277
A|Status: preliminary; translation not shown; translated from GB/EMBL/DDBJ
A|Molecule type: DNA
A|Residues: 1-16 <RES>
A|Cross-references: GB:I07899; NID:g967973; PID:g967974
R:Walgairetti, N.; Acquati, F.; Magnaghi, P.; Bruno, L.; Pontoglio, M.; Rocchi, M.; Sac
Proc. Natl. Acad. Sci. U.S.A. 89, 11584-11588, 1992
A|Title: Characterization by yeast artificial chromosome cloning of the linked apolipop
A|Reference number: A47233; MUID:93087573; EMBL:L454851

A: session: I60906
A: atus: preliminary; translated from GB/EMBL/DBDJ
A: le: type: DNA
A: lides: 1-16 <RE2>
A: ss-references: GB:M90078; NID:gl78786; PIDN:AAA35547.1; PID:G553188
A: e: apo(a) gene 1 (nomenclature of reference I52415)
A: session: A47233
A: atus: preliminary; translation not shown; translated from GB/EMBL/DBDJ
A: le: type: DNA
A: lides: 1-16 <RE5>
A: ss-references: GB:M90079; NID:gl78784; PIDN:AAA35546.1; PID:G553187
A: linoase, A.
B: linostry 31, 3113-3118, 1992
A: l: Multiple members of the plasminogen-apolipoprotein(a) gene family associated with
A: lence number: I52415; MUID:92207924; PMID:1554698
A: session: I52415
A: atus: preliminary; translated from GB/EMBL/DBDJ
A: le: type: DNA
A: lides: 1-16 <RE4>
A: ss-references: GB:M86878; NID:gl78782; PIDN:AAA51749.1; PID:G553186
A: letics:
A: e: GDB:LPA
A: ss-references: GDB:120699; OMIM:152200
A: l position: 6q26-6q27
A: e: several genes closely linked on chromosome 6 are identical in the first coding e
A: l: kringle repeats
A: efamily: apolipoprotein(a); kringle homology; trypsin homology
A: l: words: hydrolase; kringle; lipid binding; lipoprotein; trypsin homology
A: l: 19/Domains: signal sequence #status predicted <SIG>
A: l: 4548/Product: apolipoprotein(a) #status experimental <MAT>
A: l: 105/Domains: kringle homology <KR1>
A: l: 2-219/Domains: kringle homology <KR2>
A: l: 5-333/Domains: kringle homology <KR3>
A: l: 0-447/Domains: kringle homology <KR4>
A: l: 1-561/Domains: kringle homology <KR5>
A: l: 3-675/Domains: kringle homology <KR6>
A: l: 2-789/Domains: kringle homology <KR7>
A: l: 5-903/Domains: kringle homology <KR8>
A: l: 0-1017/Domains: kringle homology <KR9>
A: l: 54-1131/Domains: kringle homology <KR10>
A: l: 38-1245/Domains: kringle homology <KR11>
A: l: 32-1359/Domains: kringle homology <KR12>
A: l: 96-1473/Domains: kringle homology <KR13>
A: l: 10-1587/Domains: kringle homology <KR14>
A: l: 24-1701/Domains: kringle homology <KR15>
A: l: 38-1815/Domains: kringle homology <KR16>
A: l: 32-1929/Domains: kringle homology <KR17>
A: l: 36-2043/Domains: kringle homology <KR18>
A: l: 30-2157/Domains: kringle homology <KR19>
A: l: 94-2271/Domains: kringle homology <KR20>
A: l: 98-2385/Domains: kringle homology <KR21>
A: l: 22-2499/Domains: kringle homology <KR22>
A: l: 36-2613/Domains: kringle homology <KR23>
A: l: 50-2727/Domains: kringle homology <KR24>
A: l: 54-2841/Domains: kringle homology <KR25>
A: l: 78-2955/Domains: kringle homology <KR26>
A: l: 32-3089/Domains: kringle homology <KR27>
A: l: 36-3183/Domains: kringle homology <KR28>
A: l: 10-3297/Domains: kringle homology <KR29>
A: l: 14-3411/Domains: kringle homology <KR30>
A: l: 48-3525/Domains: kringle homology <KR31>
A: l: 52-3639/Domains: kringle homology <KR32>
A: l: 76-3753/Domains: kringle homology <KR33>
A: l: 32-3859/Domains: kringle homology <KR34>
A: l: 96-3973/Domains: kringle homology <KR35>
A: l: 10-4087/Domains: kringle homology <KR36>
A: l: 34-4201/Domains: kringle homology <KR37>

F:4228-4307/Domains: kringle homology <KR38>
F:4328-4541/Domains: trypsin homology <TRY>

Query Match 19.4%; Score 162.5; DB 1; Length 4548;
Best Local Similarity 33.1%; Pred. No. 2.4e-06;
Matches 41; Conservative 12; Mismatches 46; Indels 25; Gaps 7;

QY 27 NIHW--CN---CP-----KKFGGQHCIEIDKST---CYEGNGHFYRGKASTDTMGR 69
DB 3742 NVREYCNLTQCPVTSSVLATSTAVSEQAPTQSCPTVQDCYHGDSYRGSFSTTVTGR 3801

QY 70 PCLPWNSATVLQOQTYHAHRSDALQGLGKKNYCRNPNRRRPPWCYVQVGLKPLV--QECM 127
DB 3802 TCOSWSMT---PHWHQRTTEYYPNGGLTRNYCRNPDAEIRPWCYT---MDPSVRWEYCN 3855

QY 128 VHDC 131
DB 3856 LTQC 3859

RESULT 26
A:32869
apolipoprotein(a) (EC 3.4.21.-) - rhesus macaque (fragment)
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 22-Jun-1999
C:Accession: A32869; A30848
R:Tominson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A:Reference number: A32869; MUID:99174660; PMID:2925643
A:Accession: A32869
A:Molecule type: mRNA
A:Residues: 1-1420 <TCM>
A:Cross-references: GB:J04635; NID:G342072; PIDN:AAA36833.1; PID:G342073
C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F:50-127/Domains: kringle homology <KR1>
F:164-241/Domains: kringle homology <KR2>
F:278-355/Domains: kringle homology <KR3>
F:392-469/Domains: kringle homology <KR4>
F:506-583/Domains: kringle homology <KR5>
F:620-697/Domains: kringle homology <KR6>
F:726-803/Domains: kringle homology <KR7>
F:840-917/Domains: kringle homology <KR8>
F:954-1031/Domains: kringle homology <KR9>
F:1068-1145/Domains: kringle homology <KR10>
F:1191-1413/Domains: trypsin homology <TRY>

Query Match 19.2%; Score 161; DB 2; Length 1420;
Best Local Similarity 40.9%; Pred. No. 1.2e-06;
Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;

QY 50 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQ--QTYHAHRSDALQGLGKKNYCRNPDN 107
DB 1068 CYHNGQSYRGTFSTTVTGRTCOSWSMTPHQHKRTPENHPNDLTM-----NYCRNPDA 1122

QY 108 RRREWCYVQVGLKPLVQV--CMVHDCAD 133
DB 1123 DTGPWCFT---MDPSVRREYCNLTRCSD 1147

RESULT 27
C61545
plasmin (EC 3.4.21.7) precursor - goat (fragments)
N:Alternate names: plasminogen
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C:Accession: C61545
R:Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A:Title: Structural aspects of the plasminogen of various species.
A:Reference number: A61545; MUID:8905015; PMID:3168975
A:Accession: C61545

R.Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A>Title: Structural aspects of the plasminogen of various species.
A:Reference number: A61545; MUID:8905015; PMID:3168975
A:Accession: E61545
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-120 <SCH>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolog
C:Keywords: hydrolase; serine proteinase
F:37-114/Domain: kringle homology <K34>

Query Match 19.0%; Score 159; DB 2; Length 120;
Best Local Similarity 38.9%; Pred. No. 2.2e-07;
Matches 37; Conservative 8; Mismatches 32; Indels 18; Gaps 5;
QY 46 KSKTVEGNHGPHYRGKASTDTMGRCPLCPWNSATVLQQTYYAHRSDAL-----QLGGLGKN 100
DB 33 KVEQCYHNGNQSVRGTSSTITGRKQSWSWT-----PHRHEKTPHFPEAGL-TWN 84
QY 101 YCRNPDNRPRPCYVQVGLKPLV--QECMVHDCAD 133
DB 85 YCRNPDADKSPWCYT---TDPSPVWEFCNLKCLD 116

RESULT 30
PL50
plasmin (EC 3.4.21.7) precursor - bovine
N:Alternate names: plasminogen
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Sep-1987 #sequence revision 28-Apr-1995 #text_change 18-Jun-1999
C:Accession: S45046; A25835; I45961; S03736
R:Berglund, L.; Andersen, M.D.; Petersen, T.E.
submitted to the EMBL Data Library, May 1994
A:Description: Cloning and characterization of the bovine plasminogen cDNA.
A:Reference number: S45046
A:Accession: S45046
A:Molecule type: mRNA
A:Residues: 1-812 <BER>
A:Cross-references: EMBL:X79402; NID:G494962; PIDN:CAA55939.1; PID:G494963
A:Experimental source: liver
A:Note: It is uncertain whether Met-1 or Met-8 is the initiator
R:Schaller, J.; Moser, P.W.; Danneberg-Muller, G.A.K.; Rosselet, S.J.; Kampfer, U.; Ric
Eur. J. Biochem. 149, 267-278, 1985
A>Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plas
A:Reference number: A25835; MUID:85203306; PMID:3946532
A:Accession: A25835
A:Molecule type: protein
A:Residues: 27-334, 'D', 336-515, 'H', 517-554, 'L', 556-812 <SCH>
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A>Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A:Reference number: I45961; MUID:85023311; PMID:6148961
A:Accession: I45961
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 706-743, 'R', 745-812 <MAL>
A:Cross-references: GB:K02935; NID:G163551; PIDN:AAA30714.1; PLD:G163552
R:Brinholzer, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.
Eur. J. Biochem. 114, 465-470, 1981
A>Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A:Reference number: S03736; MUID:81212097; PMID:7238497
A:Accession: S03736
A:Molecule type: protein
A:Residues: 27-83 <BRU>
C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a v
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolog
C:Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma
F:1-26/Domain: signal sequence #status predicted <SIG>
F:18-101/Domain: plasminogen-related protein precursor homolog <PRU>

F	-812/Product: plasminogen #status experimental <PRO>
F	-103/Domain: activation peptide #status experimental <APT>
F	4-583,584-812/Product: plasmin #status experimental <NA>
F	1-583/Domain: plasmin chain A #status experimental <MA>
F	7-188/Domain: kringle homology <KR1>
F	2-269/Domain: kringle homology <KR2>
F	2-359/Domain: kringle homology <KR3>
F	1-461/Domain: kringle homology <KR4>
F	3-564/Domain: kringle homology <KR5>
F	1-812/Domain: plasmin chain B #status experimental <BCH>
F	1-805/Domain: trypsin homology <TRY>
F	80,60-68,110-188,131-171,159-183,192-269,195-323,213-252,241-264,282-359,303-342,333-342,343-344,345-346,347-348,349-350,351-352,353-354,355-356,357-358,359-360,361-362,363-364,365-366,367-368,369-370,371-372,373-374,375-376,377-378,379-380,381-382,383-384,385-386,387-388,389-390,391-392,393-394,395-396,397-398,399-400,401-402,403-404,405-406,407-408,409-410,411-412,413-414,415-416,417-418,419-420,421-422,423-424,425-426,427-428,429-430,431-432,433-434,435-436,437-438,439-440,441-442,443-444,445-446,447-448,449-450,451-452,453-454,455-456,457-458,459-460,461-462,463-464,465-466,467-468,469-470,471-472,473-474,475-476,477-478,479-480,481-482,483-484,485-486,487-488,489-490,491-492,493-494,495-496,497-498,499-500,501-502,503-504,505-506,507-508,509-510,511-512,513-514,515-516,517-518,519-520,521-522,523-524,525-526,527-528,529-530,531-532,533-534,535-536,537-538,539-540,541-542,543-544,545-546,547-548,549-550,551-552,553-554,555-556,557-558,559-560,561-562,563-564,565-566,567-568,569-570,571-572,573-574,575-576,577-578,579-580,581-582,583-584,585-586,587-588,589-590,591-592,593-594,595-596,597-598,599-600,601-602,603-604,605-606,607-608,609-610,611-612,613-614,615-616,617-618,619-620,621-622,623-624,625-626,627-628,629-630,631-632,633-634,635-636,637-638,639-640,641-642,643-644,645-646,647-648,649-650,651-652,653-654,655-656,657-658,659-660,661-662,663-664,665-666,667-668,669-670,671-672,673-674,675-676,677-678,679-680,681-682,683-684,685-686,687-688,689-690,691-692,693-694,695-696,697-698,699-700,701-702,703-704,705-706,707-708,709-710,711-712,713-714,715-716,717-718,719-720,721-722,723-724,725-726,727-728,729-730,731-732,733-734,735-736,737-738,739-740,741-742,743-744,745-746,747-748,749-750,751-752,753-754,755-756,757-758,759-760,761-762,763-764,765-766,767-768,769-770,771-772,773-774,775-776,777-778,779-780,781-782,783-784,785-786,787-788,789-790,791-792,793-794,795-796,797-798,799-800,801-802,803-804,805-806,807-808,809-810,811-812,813-814,815-816,817-818,819-820,821-822,823-824,825-826,827-828,829-830,831-832,833-834,835-836,837-838,839-840,841-842,843-844,845-846,847-848,849-850,851-852,853-854,855-856,857-858,859-860,861-862,863-864,865-866,867-868,869-870,871-872,873-874,875-876,877-878,879-880,881-882,883-884,885-886,887-888,889-890,891-892,893-894,895-896,897-898,899-900,901-902,903-904,905-906,907-908,909-910,911-912,913-914,915-916,917-918,919-920,921-922,923-924,925-926,927-928,929-930,931-932,933-934,935-936,937-938,939-940,941-942,943-944,945-946,947-948,949-950,951-952,953-954,955-956,957-958,959-960,961-962,963-964,965-966,967-968,969-970,971-972,973-974,975-976,977-978,979-980,981-982,983-984,985-986,987-988,989-990,991-992,993-994,995-996,997-998,999-1000
F	1-667,762/Active site: His, Asp, Ser #status predicted
F	try Match 19.0%; Score 159; DB 1; Length 812;
F	st Local Similarity 28.0%; Pred. No. 1.1e+06;
F	ches 49; Conservative 15; Mismatches 55; Indels 56; Gaps 11;
QY	5 HQVPSNCDCLN-----GGTCVNK7FSNIHW--CNCPPKFGGCHCEIKSKT 49
DB	315 NRTPENPFCRNLBNYCNRNPGEKAPMCYTIN--SEVRMEYCTIPS-----CESSPLST 366
QY	50 -----CYEGNGHFYRGKASTDTMGPRCLPWNSATVLQOTYHAH---RS 89
DB	367 ERMDVPVPEQTVPVQCCHGNQSRYGTSSITIGTKCOSWS-----WTPRHHLKTPE 421
QY	90 DALQLGLGKNYCRNPDRRRCPCYVOVGUKPLV--QECMVHDCADGKK--PSPS 140
DB	422 NYPNAGL-TWNYCRNPDAKSPWCYT---TDPRVRWEFCNLKCSETPEQVPAAP 472
RE	T 31
A4	32
ne	trophic receptor rorl precursor - human
N	utains: protein-tyrosine kinase (EC 2.7.1.112)
C	cies: Homo sapiens (man)
C	ce: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 19-May-2000
C	reason: A45082
P	siakowski, P.; Carroll, R.D.
J	iol. Chem. 267, 26181-26190, 1992
A	le: A novel family of cell surface receptors with tyrosine kinase-like domain.
A	erence number: A45082; MUID:93100347; PMID:1334494
A	ression: A45082
A	lecule type: mRNA
A	sidues: 1-937 <MAS>
A	ss-references: GB:M97675; NID:g337464; PIDN:AAA60275.1; PID:g337465
A	e: sequence extracted from NCBI backbone (NCBIP:120916)
C	hetics:
A	re: GDB

D 244 WCFTTDPK-RWEYCDIPCTTPPPPTVQCLGRGENYRGTVSVTSVSGTKQWRS-- 300
Q 78 TVLQOYTHAHRSDALQLGLG--KHNYCRNPDNRPRPWCYVQVGLKPLVQECMVHDCADGK 135
D 301 ---EQTPHRNTPTNPFPCNLENYCRNPDGETAPWCYT-TDSQLRWEYCEIPSCSSA 356
Q 136 KP-----SSPPEE 143
D 357 SPDQSDSSVPPEE 369

R JLT 35
A 140
P min (EC 3.4.21.7) precursor - chicken (fragment)
N ternate names: plasminogen
C Species: Gallus gallus (chicken)
C Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 16-Jul-1999
C Accession: A60140
R genes: M.; Parthy, L.
B chim. Biophys. Acta 832, 326-330, 1985
A title: The kringle 4 domain of chicken plasminogen.
A Reference number: A60140; MUID:86077796; PMID:4074753
A Accession: A60140
A molecule type: protein
A residues: 1-89 <GV>
A Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F -83/Domain: kringle homology <KR>
F -83/Binding site: carbohydrate (Asn) (covalent) #status experimental
Query Match 18.4%; Score 154; DB 2; Length 89;
Best Local Similarity 37.4%; Pred. No. 4.8e-07;
Matches 34; Conservative 9; Mismatches 36; Indels 12; Gaps 4;

Q 47 SKTCYEGNHFRGKASTDTMGRPCLPWNSATVLOQYTHAHRSDALQLGLG--KHNYCRN 104
D 3 TEECYQNGSVYRGTSFTITGKCKQAWNS-----MSPHRNKTESHPNADLRQNYCRN 57
Q 105 PDNRPRPWCYVQVGLKPLV--QECMVHDCAD 133
D 58 PDNRSPWCYT---TDPVSRWEYCNLKRCS 85

R JLT 36
I 260
P min (EC 3.4.21.7) precursor - western European hedgehog
C Species: Erinaceus europaeus (western European hedgehog)
C Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
C Accession: I46260
R wn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong
Biol. Chem. 270, 24004-24009, 1995
A title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprotein
A Reference number: I46259; MUID:96025778; PMID:7592597
A Accession: I46260
A Status: Preliminary; translated from GB/EMBL/DBJ
A molecule type: mRNA
A residues: 1-810 <LAW>
A Cross-references: EMBL:U33171; NID:g1046360; PID:g1046361
C Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C Keywords: hydrolase; serine proteinase
F -86/Domain: plasminogen-related protein precursor homology <PLPH>
F -13-181/Domain: kringle homology <KR1>
F -35-262/Domain: kringle homology <KR2>
F -75-352/Domain: kringle homology <KR3>
F -79-456/Domain: kringle homology <KR4>
F -12-561/Domain: kringle homology <KR5>
F -12-803/Domain: trypsin homology <TR>

Query Match 18.3%; Score 153.5; DB 2; Length 810;
Best Local Similarity 27.3%; Pred. No. 3.5e-06;
Matches 39; Conservative 12; Mismatches 47; Indels 45; Gaps 6;

QY 6 QVPSNCDCLN-----GGTCVNKYFSNIHW--CNCPKKFGGQHCIDSKT- 49
DB 309 RTENYPCNLDENYCNPDGEPAPWCFTN--SSVRWEPCKIP-----DCVSSASETE 360
QY 50 -----CYEGNHFRGKASTDTMGRPCLPWNSATVLOQYTHAHRSDA 91
DB 361 HSDAPVIVPEQTFFVQECYQNGQYRGTSFTITGKCKQAWNSMRPHRSKTPENYPD 420
QY 92 LQLGLGHNYCRNPDNRPRPWCY 114
DB 421 ADLTM---NYCRNPDGDKGWCY 440

RESULT 37
A40522
P min (EC 3.4.21.7) precursor - rat (fragment)
C Species: Rattus norvegicus (Norway rat)
C Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
C Accession: A40522
R Kanakas, J.J.; Makker, S.P.
J. Biol. Chem. 266, 10825-10829, 1991
A title: Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor
A Reference number: A40522; MUID:91250378; PMID:1645711
A Accession: A40522
A Status: Preliminary
A molecule type: mRNA
A residues: 1-169 <KAN>
A Cross-references: GB:M62832; NID:9206215; PID:AAA41884.1; PID:G544488
A Note: the authors translated the codon TGT for residue 76 as Ala
C Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F -34-112/Domain: kringle homology <KR>
F -34-112,55-95,83-107/Disulfide bonds: #status predicted
Query Match 18.2%; Score 152.5; DB 2; Length 169;
Best Local Similarity 30.2%; Pred. No. 1.1e-06;
Matches 42; Conservative 18; Mismatches 32; Indels 47; Gaps 9;

QY 40 QHCEI-----DKS-----KTCYEGNHFRGKASTDTMGRPCLPWNSAT 78
DB 3 EYCEIPSCGSSVSPDQSDSSVLPEQTFVQECYQNGKSYRGTSFTITGKCKQSW---- 58
QY 79 VLOQYTHAHRSDALQL---GLGKHNYCRNPDN---RRPWCYVQVGLKPLV--QECMVHDC 132
DB 59 -VSMTPHSHSKTPANFPDSGL-EMNYCRNPDNDQRPWCFT---TDPVSRWEYCNLKRCS 113
QY 133 D-----GKPPSP 140
DB 114 ETGCGVAESAIVPQVPSAP 132

RESULT 38
A47136
P macrophage-stimulating protein 1 precursor - human
C Species: Homo sapiens (man)
C Date: 03-May-1994 #sequence_revision 14-Nov-1997 #text_change 18-Jun-1999
C Accession: A40331; A47136; A61395
R Han, S.; Stuart, L.A.; Degen, S.J.F.
Biochemistry 30, 9768-9780, 1991
A title: Characterization of the DNFI582 locus on human chromosome 3: identification of
A Reference number: A40331; MUID:92002016; PMID:1655021
A Accession: A40331
A molecule type: DNA
A residues: 1-711 <HAL>
A Cross-references: GB:M74179
A Accession: B40331
A molecule type: mRNA
A residues: 1-711 <HA2>
A Cross-references: GB:M74178; NID:gl83976; PID:AAAS0165.1; PID:gl83977
R Yoshimura, T.; Yuhki, N.; Wang, M.H.; Skeel, A.; Leonard, E.J.
J. Biol. Chem. 268, 15461-15468, 1993
A title: Cloning, sequencing, and expression of human macrophage stimulating protein (M

C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase; PRO>
F;1-33,34-117,118-455/Product: plasminogen (fragments) #status experimental <PRO>
F;1-33/Domain: activation peptide (fragment) #status experimental <APT>
F;34-117,118-225,226-455/Product: plasmin (fragments) #status experimental <MAT>
F;37-114/Domain: kringle homology <KR4>
F;118-455/Product: miniplasminogen #status experimental <MIN>
F;126-205/Domain: kringle homology <KR5>
F;226-455/Domain: plasmin chain B #status experimental <BCH>
F;226-448/Domain: trypsin homology <TRY>
F;267,310,405/Active site: His, Asp, Ser #status predicted

Query Match 17.8%; Score 149; DB 2; Length 455;
Best Local Similarity 34.0%; Pred. No. 5.4e-06;
Matches 36; Conservative 17; Mismatches 39; Indels 14; Gaps 6;

Oy 40 QHCEIDSKTCYEGNGHFYRGKASDTDMGRPCLPWNSATV--LQOITYHAHRSDALQLGLG 97
 :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db 28 EECEA-KVQDCYQDKGESYRGTSITVTGKKCSWSSMTPHHQXTPEKYPNADLTJ--- 83

Oy 98 KHNVCNPDNRPRPCYVOVGKLPLY-QECMVHCADG-KKPSPS 140
 ::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db 84 --NYCNPDGDKGPCYT---TDPVVRWFECNLKCKSETVQEPSEP 124

RESULT 40
B45082
neurotrophic receptor ror2 precursor - human
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 19-May-2000
C;Accession: B45082
R;Maslowski, P.; Carroll, R.D.
J. Biol. Chem. 267, 26181-26190, 1992
A;Title: A novel family of cell surface receptors with tyrosine kinase-like domain.
A;Reference number: A45082; MUID:93100347; PMID:1334494
A;Accession: B45082
A;Molecule type: mRNA
A;Residues: 1-943 <MAS>
A;Cross-references: GB:M97639; NID:G337466; PIDN:AAA60276.1; PID:G337467
A;Note: sequence extracted from NCBI backbone (NCBIP:120918)
C;Genetics:
A;Gene: GDB:NTRK32
A;Cross-references: GDB:136454
A;Map position: 5p21-6p21
C;Superfamily: neurotrophic receptor ror; immunoglobulin homology; kringle homology; pro-
C;Keywords: APP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyro-
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-937/Product: neurotrophic receptor ror2 #status predicted <MAT>
F;78-137/Domain: immunoglobulin homology <IMM>
F;316-394/Domain: kringle homology <KRG>
F;412-428/Domain: transmembrane #status predicted <TMN>
F;471-753/Domain: protein kinase homology <KIN>
F;479-487/Region: protein kinase ATP-binding motif
F;70,188,318/Binding site: carbohydrate (Asn) #status predicted

Query Match 17.7%; Score 148; DB 2; Length 943;
Best Local Similarity 34.3%; Pred. No. 1.2e-05;
Matches 37; Conservative 12; Mismatches 41; Indels 18; Gaps 6;

Oy 32 NC-----PKKGFGQCEIDKSKTCYEGNGHFYRGKASDTDMGRPCLPWNSATVLQQTYY 85
 :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db 299 NCWRIGIPAERLGRYH-----QCYNCGMDYRGTAFTKSGHQCPW--ALQHPHS 349

Oy 86 AHRSDALQLGLGHKNYCNPDRNR--PWCYVQVGLKPLVQECMVHCA 132
 :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db 350 LSTDFPFELG-GGHAYCRNPGQMEGPWCFTQ-NKNVRMELCDVPSCS 395

RESULT 41
A40332
macrophage-stimulating protein 1 precursor - mouse
N;Alternate names: hepatocyte growth factor-like protein
C;Species: Mus musculus (house mouse)

ate: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 18-Jun-1999
C ccession: A40332; B40332
F egen, S.J.F.; Stuart, L.A.; Han, S.; Jamison, C.S.
E Chemistry 30, 9781-9791, 1991
F title: Characterization of the mouse cDNA and gene coding for a hepatocyte growth factor
F reference number: A40332; MUID:92002017; PMID:1832957
A ccession: A40332
A olecule type: DNA
A esidues: 1-716 <DEG>
A ross-references: GB:M74180; NID:G193831; PIDN:AAA50166.1; PID:G193832
F ccession: B40332
A olecule type: mRNA
A esidues: 1-18, 'P', 20-716 <DEG2>
A ross-references: GB:M74181; NID:G193833; PIDN:AAA50167.1; PID:G193834
C etics:
A etons: 18/1; 67/2; 105/1; 143/2; 189/1; 229/2; 269/1; 334/2; 378/1; 412/2; 459/1; 47
C omplex: disulfide-bonded heterodimer of chains derived from the same precursor
C uperfamily: hepatocyte growth factor; kringie homology; trypsin homology
C ywords: duplication; glycoprotein; growth factor; kringie
F 31/Domain: signal sequence #status predicted <SIG>
F 9-488,489-716/Product: macrophage-stimulating protein 1 #status experimental <MAT>
F 9-483/Domain: alpha chain #status experimental <ACH>
F 10-186/Domain: kringie homology <KR1>
F 91-268/Domain: kringie homology <KR2>
F 92-370/Domain: kringie homology <KR3>
F 79-457/Domain: kringie homology <KR4>
F 94-711/Domain: beta chain #status experimental <BCH>
F 39-709/Domain: trypsin homology <TRY>
F 2,173,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 17.6%; Score 147.5; DB 1; Length 716;
Best Local Similarity 36.5%; Pred. No. 1.1e-05;
atches 38; Conservative 11; Mismatches 34; Indels 21; Gaps 7;
C 50 CYEGNGHFYRGKASTDTMGPCPLPNSATVLQTYHAHR-----SDALQLGLGKENVCRN 104
D 379 CYHSGEQRGVSKTRKGVQCOHWSSET-----PHRQPTPTSAPOAGL-EANFCRN 430
C 105 PD-NRRRPMCYVQGLKP--LVQECMVHDCADGKFS--SPPEE 143
D 431 PDGDSHGPMWCYT--LDPDILPDYCALQRCDDQPPSILDPDQ 471
F ULT 42
F 318
A lipoprotein(a) - western European hedgehog (fragment)
C ectes: Erinaceus europaeus (western European hedgehog)
C ate: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
C ccession: T18518
F awn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong
F iol. Chem. 270, 24004-24009, 1995
A title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprotein
A reference number: I46259; MUID:96025778; PMID:7592597
A ccession: T18518
A tatus: preliminary; translated from GB/EMBL/DBJ
A olecule type: mRNA
A esidues: 1-2869 <LAW>
A ross-references: EMBL:U33170; NID:G1046358; PID:G1046359; PIDN:AA48522.1
A xperimental source: liver
A oment: The lipoprotein Lp(a), a major inherited risk factor for atherosclerosis, con
e apolipoprotein(a).
Query Match 17.6%; Score 147; DB 2; Length 2869;
Best Local Similarity 29.3%; Pred. No. 4e-05;
atches 39; Conservative 11; Mismatches 61; Indels 22; Gaps 6;
C 19 CVSNKYFSNIHWCNCFKFGQHCIDKSKT-----CYEGNGHFYRGKASTDTMGPR 70
D 2558 CYTNSAMRWYCISFA-----CESPTPTTEHLVPEQCLEGNGENYQGNMAITVSGOP 2611
C 71 CLPNSATVLQTYHAHRSDALQLGLGKENVCRNPDNRRPWCYVQVGLKPLVQCMVHD 130
I 2612 CQGRKQTPHRRHYETENVPFSKNL-FG--NYCRNPDGEIAPWCYT-TNSAVRWYCISPT 2667

RESULT 43

T18840 hypothetical protein C01G6.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T18840

R:Barks, M.

submitted to the EMBL Data Library, August 1994

A:Reference number: Z19029

A:Accession: T18840

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-806 <WIL>

A:Cross-references: EMBL:Z35595; PIDN:CAA84639.1; GSPDB:GN00020; CESP:C01G6.8

A:Experimental source: clone C01G6

C:Genetics:

A:Gene: CESP:C01G6.8

A:Map position: 2

A:Introns: 36/3; 170/3; 217/3; 636/3; 760/3

Query Match 17.5%; Score 146.5; DB 2; Length 806;

Best Local Similarity 25.1%; Pred. No. 1.5e-05;

Matches 43; Conservative 25; Mismatches 60; Indels 43; Gaps 9;

RESULT 44

JC5061

macrophage-stimulating protein 1 precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Jun-2000

C:Accession: JC5061

R:Ohnishi, K.; Iwama, A.; Matsuno, K.; Ezaki, T.; Sakamoto, O.; Hamaguchi, I.; Takasu,

Biochem. Biophys. Res. Commun. 227, 273-280, 1996

A:Title: Molecular cloning of rat macrophage-stimulating protein and its involvement in

A:Reference number: JC5061; MUID:97011126; PMID:8858136

A:Accession: JC5061

A:Molecule type: mRNA

A:Residues: 1-716 <OHS>

A:Cross-references: EMBL:X95096; NID:G1669718; PIDN:CAA64473.1; PID:G1669719

C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor

C:Superfamily: hepatocyte growth factor; kringie homology; trypsin homology

C:Keywords: duplication; glycoprotein; growth factor; kringie

F:1-31/Domain: signal sequence #status predicted <SIG>

F:32-488,489-716/Product: macrophage-stimulating protein 1 #status predicted <MAT>

F:32-488,489-716/Product: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>

F:191-268/Domain: kringie homology <KR12>

F:292-370/Domain: kringie homology <KR13>

F:379-457/Domain: kringie homology <KR14>

F:489-709/Domain: macrophage-stimulating protein 1 beta chain #status predicted <BCH>

F:72,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.4%; Score 145.5; DB 1; Length 716;

Best Local Similarity 33.6%; Pred. No. 1.7e-05;

30 WC---NCPKFGGCHCEIDKSKTCYEG-----NGHFYRGKASTDTWGRPCLPWSATVLIQ 81
 168 WCVTTRESVRF--QSGI---KSCRAVCVWNGEDYRGVEDVTSRGRCQWDLOPHES 222
 82 QYHAHRSALQGLGKHNYCRNRPWPCVQVGLKPLVQES--CMVHDCADGKPKSS 139
 223 HPFHPKFPDKAL---KDNVCRNPDASERFWCYT---TDPNVEREFCDLPSG-----PVL 272
 140 PP 141
 273 PP 274
 JT 45
 he bocyte growth factor precursor [validated] - human
 N; ternate names: hepatopoietin A; scatter factor
 C; scies: Homo sapiens (man)
 C; te: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 08-Dec-2000
 C; session: JH0579; JH0333; A41140; B36677; A36677; A33512; A39006; PH0114; A37796; S08
 R; ki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
 Ge 102, 213-219, 1991
 A; ele: Organization of the human hepatocyte growth factor-encoding gene.
 A; Reference number: JH0579; MUID:91340155; PMID:1831432
 A; session: JH0579
 A; lecule type: DNA
 A; sidues: 1-728 <SEK>
 A; oss-references: DDBJ:D90318
 A; te: the authors translated the codon GAA for residue 662 as Gly
 R; ki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
 su litted to JIPID, March 1991
 A; scription: Organization of the human hepatocyte growth factor-encoding gene.
 A; Reference number: JH0579
 A; session: JH0579
 A; lecule type: DNA
 A; sidues: 1-481, 'ET', 484-728 <SE2>
 R; idner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder, H.
 Pr . Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991
 A; ele: Evidence for the identity of human scatter factor and human hepatocyte growth factor
 A; Reference number: A41140; MUID:91334393; PMID:1831466
 A; session: A41140
 A; lecule type: mRNA
 A; sidues: 1-728 <WEI>
 A; oss-references: GB:M73239; NID:9337935; PIDN:AAA4239.1; PID:G337936
 R; ki, T.; Ihara, I.; Sugimura, A.; Shimonishi, M.; Nishizawa, T.; Asami, O.; Hagiya, M.
 Bi hem. Biophys. Res. Commun. 172, 321-327, 1990
 A; ele: Isolation and expression of cDNA for different forms of hepatocyte growth factor
 A; Reference number: A36677; MUID:91025062; PMID:2145836
 A; session: B36677
 A; lecule type: mRNA
 A; sidues: 1-728 <SE3>
 A; oss-references: GB:M60718; NID:9184031; PIDN:AAA52648.1; PID:G184032
 A; session: A36677
 A; lecule type: mRNA
 A; sidues: 1-161, 167-728 <SE4>
 A; oss-references: EMBL:X16323
 A; perimental source: leukocyte
 R; yazawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nakaya
 Bi hem. Biophys. Res. Commun. 163, 967-973, 1989
 A; ele: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth factor
 A; Reference number: A33512; MUID:8932017; PMID:2528952
 A; session: A33512
 A; tus: not compared with conceptual translation
 A; sidues: 1-728 <MIY>
 A; oss-references: GB:M29145; NID:9184041; PIDN:AAA52650.1; PID:G306846
 R; bin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.; Hix
 Pr . Natl. Acad. Sci. U.S.A. 88, 415-419, 1991
 A; ele: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepatocyte
 A; Reference number: A39006; MUID:9110540; PMID:1824873
 A; session: A39006

A; Molecule type: mRNA
 A; Residues: 1-161, 167-728 <RUB>
 A; Cross-references: GB:M55379
 A; Experimental source: embryonic lung
 R; Yoshizawa, Y.; Arakaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayama, I.
 Biochem. Biophys. Res. Commun. 175, 660-667, 1991
 A; Title: Identification of the N-terminal residue of the heavy chain of both native and
 A; Reference number: PH0114; MUID:91207365; PMID:1826837
 A; Accession: PH0114
 A; Molecule type: protein
 A; Residues: 32-43, 53-58 <YOS>
 A; Experimental source: plasma
 R; Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier, W.
 J. Cell Biol. 111, 2097-2108, 1990
 A; Title: Scatter factor: molecular characteristics and effect on the invasiveness of epi
 A; Reference number: A37796; MUID:91035621; PMID:2146276
 A; Accession: A37796
 A; Molecule type: protein
 A; Residues: 86-91, 329-344, 356-363 'XX', 366-370, 425-434, 442-447, 'X', 449-450, 543-546, 'X', 5
 R; Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimonishi, M.; Sugimura, A.; Tash
 Nature 342, 440-443, 1989
 A; Title: Molecular cloning and expression of human hepatocyte growth factor.
 A; Reference number: S06794; MUID:90066676; PMID:2531289
 A; Accession: S06794
 A; Molecule type: mRNA
 A; Residues: 1-31, 'HK', 34-77, 'N', 79-292, 'V', 294-299, 'W', 301-316, 'A', 318-335, 'K', 337-386, '
 A; Cross-references: EMBL:X16323; NID:932081; PIDN:CAA34387.1; PID:G32082
 A; Experimental source: liver
 A; Note: the authors translated the codon CAG for residue 727 as Glu
 A; Note: part of this sequence, including the amino end of both the alpha and beta chains
 R; Hartmann, G.; Naldini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Birchm
 Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992
 A; Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth fact
 A; Reference number: I59214; MUID:93087571; PMID:1280830
 A; Accession: I59214
 A; Status: preliminary; translated from GB/EMBL/DBJ
 A; Molecule type: mRNA
 A; Residues: 1-288, 'ET' <HAR>
 A; Cross-references: GB:I02931; NID:G184033; PIDN:AAA52649.1; PID:G184034
 R; Miyazawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.
 Eur. J. Biochem. 197, 15-22, 1991
 A; Title: An alternatively processed mRNA generated from human hepatocyte growth factor 9
 A; Reference number: S15443; MUID:91200041; PMID:1826653
 A; Accession: S15443
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 1-288, 'ET' <MIY>
 A; Cross-references: EMBL:X57574; NID:932083; PIDN:CAA40802.1; PID:G32084
 R; Shima, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Higashio, K.
 Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991
 A; Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblasts: cloning
 A; Reference number: I52253; MUID:92062058; PMID:1835383
 A; Accession: I52253
 A; Status: preliminary; translated from GB/EMBL/DBJ
 A; Molecule type: mRNA
 A; Residues: 161-166 <SHI>
 A; Cross-references: GB:S62561; NID:9237996; PIDN:AAB20169.1; PID:G237997
 C; Genetics:
 A; Gene: GDB:HGF
 A; Cross-references: GDB:127524; OMIM:142409
 A; Map position: 7q21.1-7q21.1
 A; Introns: 30/1; 85/2; 123/1; 161/2; 209/1; 249/1; 289/1; 347/2; 390/1; 424/2; 469/1; 48
 C; Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C; Function:
 A; Description: stimulates mitosis of hepatocytes and other cells
 A; Note: does not have proteinase activity
 C; Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
 C; Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyr
 F; 1-31/Domain: signal sequence #status predicted <SIG>
 F; 32-494, 495-728/Product: hepatocyte growth factor #status experimental <WAT>
 F; 32-494/Domain: alpha chain #status experimental <ACH>
 F; 128-206/Domain: kringle homology <KR1>
 F; 211-288/Domain: kringle homology <KR2>

F 05-383/Domain: kringle homology <KR3>
 F 91-469/Domain: kringle homology <KR4>
 F 95-728/Domain: beta chain #status experimental <BCH>
 F 95-716/Domain: trypsin homology <TRY>
 F 2/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
 F 34,402,566,653/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F 87-604/Disulfide bonds: #status predicted
 Query Match 17.3%; Score 144.5; DB 1; Length 728;
 Best Local Similarity 29.6%; Pred. No. 2.1e-05;
 Matches 45; Conservative 13; Mismatches 69; Indels 25; Gaps 10;
 C 2 NELHQPNSCDD--LNGGTCVSNKYFSNIHWC-----NCPKKFGGQ--HCEIDKSKTCYE 52
 D 335 HEHDMTPENFKCKDLRENYC-RNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQCYR 393
 Q 53 GNGHYRGKASTDTMGRCPLPNSATVLTQYTHAH---RSDALQLGLGKHNYCRNP-DNR 108
 D 394 GNGKRYMGNLSQTRSGLTCSMWDXN---MEDLHRHIFWEPDASKL---NENYCRNPDDDA 447
 Q 109 RPPWCVVQVGLKPLV--QECMVHDCADGKKPS 138
 D 448 HGPWCYTG---NPLVPWDYCPISRCEGDTTPT 476
 R ULT 46
 I 285
 h atocyte growth factor/scatter factor - chicken (fragment)
 C species: Gallus gallus (chicken)
 C Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
 C Accession: I51285
 R Kretz, A.; Stern, C.D.; Thery, C.; Ireland, G.W.; Aparicio, S.; Sharpe, M.J.; Gherard
 D Alignment 121, 813-824, 1995
 A Title: A role for HGF/SF in neural induction and its expression in Hensen's node durin
 A Reference number: I51285; MUID:95237013; PMID:7720585
 A Accession: I51285
 A Status: preliminary; translated from GB/EMBL/DBJ
 A Molecule type: mRNA
 A Residues: 1-411 <STR>
 A Cross-references: GB:S77480; NID:9998675; PID:9998676
 C Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
 F 24-197/Domain: kringle homology <KR2>
 F 32-279/Domain: kringle homology <KR3>
 F 36-374/Domain: kringle homology <KR3>
 Query Match 17.0%; Score 142; DB 2; Length 411;
 Best Local Similarity 28.4%; Pred. No. 2.1e-05;
 Matches 38; Conservative 18; Mismatches 48; Indels 30; Gaps 7;
 C 30 WC-----NCPKKFGG-QHCEID-----KSKTCYEGNGHYRGKASTDTMGRCPLPW 74
 D 261 WCYTLDPNTPWEFCAIKCDVGLNSTEAVAEITTCIQGQGGYRGTVNTWSGLQCQRW 320
 Q 75 NSATVLQQTYHAHR--SDALQLGLGKHNYCRNPDRRPPWCY-----VQVGLKPLVQECM 127
 D 321 DS-----QPPHQHNTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPKCD 375
 Q 128 V---HDCADGKKPS 138
 D 376 VSNEDQCYRGNGKS 389
 R JLT 47
 A 344
 h atocyte growth factor precursor - rat
 N Alternate names: hepatoin A; scatter factor
 C species: Rattus norvegicus (Norway rat)
 C Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 21-Jul-2000
 C Accession: A35644; S13211
 R Ashiro, K.; Hagiya, M.; Nishizawa, T.; Seki, T.; Shimonishi, M.; Shimizu, S.; Nakamura
 P Natl. Acad. Sci. U.S.A. 87, 3200-3204, 1990
 A Title: Deduced primary structure of rat hepatocyte growth factor and expression of the
 A Reference number: A35644; MUID:90222197; PMID:2139229

A Accession: A35644
 A Status: preliminary
 A Molecule type: mRNA
 A Residues: 1-728 <TAS>
 A Cross-references: GB:D90102; GB:M32987; NID:9220766; PIDN:BAAL4133.1; PID:9220767
 A Note: the authors translated the codon GAG for residue 70 as Gln, GAC for residue 417
 R Okajima, A.; Miyazawa, K.; Kitamura, N.
 Eur. J. Biochem. 193, 375-381, 1990
 A Title: Primary structure of rat hepatocyte growth factor and induction of its mRNA du
 A Reference number: S13211; MUID:91031482; PMID:2146117
 A Accession: S13211
 A Status: preliminary
 A Molecule type: mRNA
 A Residues: 1-728 <CA>
 A Cross-references: EMBL:X54400; NID:956353; PIDN:CAA38266.1; PID:94539554
 C Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C Function:
 A Description: stimulates mitosis of hepatocytes and other cells
 A Note: does not have proteinase activity
 C Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
 C Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; py
 F 1-32/Domain: signal sequence #status predicted <SIG>
 F 56-495/Product: hepatocyte growth factor #status predicted <MAT>
 F 56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
 F 129-207/Domain: kringle homology <KR1>
 F 212-289/Domain: kringle homology <KR2>
 F 306-384/Domain: kringle homology <KR3>
 F 392-470/Domain: kringle homology <KR4>
 F 496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
 F 496-728/Domain: trypsin homology <TRY>
 F 33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
 F 295,403,569,656/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F 488-607/Disulfide bonds: #status predicted
 Query Match 17.0%; Score 142; DB 1; Length 728;
 Best Local Similarity 28.8%; Pred. No. 3.5e-05;
 Matches 45; Conservative 11; Mismatches 62; Indels 38; Gaps 11;
 QY 5 HOV-PSNCDG--LNGGTCVSNKYFSNIHWC-----NCPKKFGGQHCIEDKSK 48
 DB 338 HDITPENFKCKDLRENYC-RNPDGABSPWCFTTDPNIRVGYCSQIPK-----CDVSSGQ 390
 QY 49 TCYEGNGHYRGKASTDTMGRCPLPNSATVLTQYTHAH---RSDALQLGLGKHNYCRNP 105
 DB 391 DCYRGNGKRYMGNLSQTRSGLTCSMWDXN---MEDLHRHIFWEPDASKL---TKNYCRNP 444
 QY 106 -DNRPRPWCVVQVGLKPLV--QECMVHDCADGKKPS 138
 DB 445 DDDAHGWCYTG---NPLVPWDYCPISRCEGDTTPT 477
 RESULT 48
 A60185
 hepatocyte growth factor precursor - mouse
 N Alternate names: hepatoin A; scatter factor
 C Species: Mus musculus (house mouse)
 C Date: 03-Mar-1993 #sequence_revision 26-May-1994 #text_change 16-Jun-2000
 C Accession: J02117; PC2064; A60185; S43416; S45521; S1173; S10966; 148758; J00231
 R Sasaki, M.; Nishio, M.; Sasaki, T.; Enami, J.
 Biochem. Biophys. Res. Commun. 199, 772-778, 1994
 A Title: Identification of mouse mammary fibroblast-derived mammary growth factor as he
 A Reference number: J02117; MUID:94183257; PMID:8135822
 A Accession: J02117
 A Molecule type: mRNA
 A Residues: 1-728 <SAS2>
 A Cross-references: GB:D10212; NID:9220435; PIDN:BAAL01064.1; PID:9220436
 A Experimental source: fibroblast, COS-1 cell
 A Note: submitted to JIPID, May 1993
 A Accession: PC2064
 A Molecule type: protein
 A Residues: 456-504 <SAS2>
 R Rosen, E.M.; Meromsky, L.; Setter, E.; Vinter, D.W.; Goldberg, I.D.
 Proc. Soc. Exp. Biol. Med. 195, 34-43, 1990

A: title: Purified scatter factor stimulates epithelial and vascular endothelial cell mig
 A: Reference number: A60185; MUID:90377927; PMID:2144630
 A: session: A60185
 A: lecul type: protein
 A: sidues: 'X',184-188,'KX',191-192,'X',194,'XX',197;357-364,'XX',367;375-377,'E',379,'
 R: Y.; Michalopoulos, G.K.; Zarnegar, R.
 R: im. Biophys. Acta 1216, 299-303, 1993
 A: title: Molecular cloning and characterization of cDNA encoding mouse hepatocyte growth
 A: Reference number: S43416; MUID:94060105; PMID:8241272
 A: session: S43416
 A: utus: preliminary
 A: lecul type: mRNA
 A: sidues: 1-728 <LIU>
 A: sse-references: EMBL:X72307
 R: Y
 A: tted to the EMBL Data Library, May 1993
 A: Reference number: S45521
 A: session: S45521
 A: utus: preliminary
 A: lecul type: mRNA
 A: sidues: 1-563,'H',565-728 <LI2>
 A: sse-references: EMBL:X72307
 R: A.; Fellows, J.; Young, S.; Pappin, D.; Rahman, D.
 R: em. J. 278, 35-41, 1991
 A: title: Purification and Characterization of biologically active scatter factor from ra
 A: Reference number: S17173; MUID:91354223; PMID:1831975
 A: session: S17173
 A: lecul type: protein
 A: sidues: 496-517,'T',519 <COP>
 R: ardi, E.; Stoker, M.
 R: e 346, 228, 1990
 A: title: Hepatocytes and scatter factor.
 A: Reference number: S10966; MUID:90326152; PMID:2142751
 A: session: S10966
 A: utus: preliminary
 A: lecul type: protein
 A: sidues: 496-507,'X',509-512,'L',514-516,'X',518-519 <NAT>
 R: schke-Schluter, A.; Behrens, J.; Gherardi, E.; Birchmeier, W.
 R: ol. Chem. 270, 830-836, 1995
 A: title: Characterization of the scatter factor/hepatocyte growth factor gene promoter.
 A: Reference number: I48758; MUID:95122532; PMID:7822318
 A: session: I48758
 A: utus: preliminary; translated from GB/EMBL/DBJ
 A: lecul type: DNA
 A: sidues: 1-30 <RES>
 A: sse-references: EMBL:X81630; MID:G673451; PIDN:CAA57286.1; PID:G673452
 A: C: plex: disulfide-bonded heterodimer of chains derived from the same precursor
 A: C: action:
 A: e: description: stimulates mitosis of hepatocytes and other cells
 A: e: does not have proteinase activity
 A: C: e: family: hepatocyte growth factor; kringle homology; trypsin homology
 A: C: words: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyr
 A: C: 2/Domain: signal sequence #status predicted <SIG>
 A: F: 495,496-728/Product: hepatocyte growth factor #status predicted <NAT>
 A: F: 9-207/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
 A: F: 2-289/Domain: kringle homology <KR1>
 A: F: 5-384/Domain: kringle homology <KR2>
 A: F: 6-470/Domain: kringle homology <KR3>
 A: F: 5-728/Domain: kringle homology <KR4>
 A: F: 6-719/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
 A: F: 3-403,569,656/Binding site: Pyroglutamate carboxylic acid (Gln) (in mature form) #status predicted
 A: F: 1-607/Disulfide bonds: #status predicted
 A: ary Match
 A: at Local Similarity 17.0%; Score 142; DB 1; Length 728;
 A: ches 44; Conservative 12; Mismatches 62; Indels 38; Gaps 11;
 Qy 5 HQV-PSNDC--LNGGTCVSNKVFSTHWC-----NCPKKFGQHCEIDKSK 48
 Db 338 HDITPENFKCKDLRENYC-RNPDAQESPWCTTDPNIRGVCSOIPK-----CDVSSGO 390

```

Qy 49 TCYEGNGHYRKGASTDTWGRCLPWN SATVLQQTYHAH---RSDALQGLGKHNYCRNP 105
Db 391 DCYVRGNGKYNMGLSKTRSGLTCSWMDKN---MEDLHRHIFWEPDASKL---NKNYCRNP 444
Qy 106 -DNRREPWCYVQVGLKPLV--QECVHDCADCKPS 138
Db 445 DDAHGCPWYTG---NPLIPWYCFISRCGDTTFT 477
RESULT 49
I51283
hepatocyte growth factor precursor - clawed frog
N/Alternate names: hepatoinetin A, scatter factor
C/Species: Xenopus sp. (Clawed frog)
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C/Accession: I51283
R/Nakamura, H.; Tashiro, K.; Nakamura, T.; Shiokawa, K.
Mech. Dev. 49, 123-131, 1995
A/Title: Molecular cloning of Xenopus HGF cDNA and its expression studies in Xenopus ear
A/Reference number: I51283, MUID:95267690; PMID:7748783
A/Accession: I51283
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-710 <NAK>
A/Cross-references: GB:S77422; NID:G989932; PIDN:AA834354.1; PID:G989933
A/Note: the authors' translation for residue 458 (Thr) is inconsistent with the nucleotide
C/Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C/Function:
A/Description: stimulates mitosis of hepatocytes and other cells
A/Note: does not have proteinase activity
A/C/Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C/Keywords: duplication; glycoprotein; growth factor; heterodimer; kringle
F:42-477/709/Product: hepatocyte growth factor #status predicted <MAT>
F:42-477/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F:115-193/Domain: kringle homology <KR1>
F:138-275/Domain: kringle homology <KR2>
F:289-367/Domain: kringle homology <KR3>
F:375-453/Domain: kringle homology <KR4>
F:478-709/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
F:478-700/Domain: trypsin homology <TRY>
F:52,128,281,322,379,550,637,666/Binding site: carbohydrate (Asn) (covalent) #status pre
F:470-588/Disulfide bonds: #status predicted
Query Match 16.8%; Score 141; DB 1; Length 710;
Best Local Similarity 29.2%; Pred. No. 4.2e-05;
Matches 42; Conservative 12; Mismatches 64; Indels 26; Gaps 9;
Qy 4 LHQ-VPSNDC--LNGGTGVSNKYSNIHWC-----NCPKFGCGHCEIDSKTC 50
Db 320 LHNFTENYKCDLSENYC-RNPDGSESPWCTTDPNIRHCSQI---KKCOASNQOEC 375
Qy 51 YEGNGHYRKGASTDTWGRCLPWN SATVLQQTYHAHRS DALQGLGKHNYCRNP DN-RR 109
Db 376 YVNGSTYKGLTSRTRFLPCSWWEKNL---QDLKHTFNEPDVSLQKNYCRNP DNDAH 432
Qy 110 RPNCYVQVGLKPLV--QECVHDC 131
Db 433 GPMCYTD---DPFVPWDYCPISRC 453
RESULT 50
S33879
Plasmin precursor - lamprey (fragments)
N/Alternate names: plasminogen
C/Species: Petromyzontidae gen. sp. (lamprey)
C/Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 07-Nov-1997
C/Accession: S33879
R/Affolter, M.; Schaller, J.; Rickli, E.E.
Protein Seq. Data Anal. 5, 207-211, 1993
A/Title: Isolation, characterization and partial amino acid sequence of lamprey plasminogen
A/Reference number: S33879
A/Accession: S33879
A/Status: preliminary

```

Molecule type: protein
A residues: 1-15;16-34;35-44;45-59;60-76;77-111;111-138;139-158;159-178;179-216;217-236;
C superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
F 1-146/Domain: kringle homology <RR3>

Query Match 16.2%; Score 136; DB 2; Length 336;
Best Local Similarity 30.7%; Pred. No. 6.2e-05;
Matches 31; Conservative 7; Mismatches 41; Indels 22; Gaps 3;

Q 34 PKKFGQHCEIDKSKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQ 93
D 71 PSEFAG-----LTTACVKGTEGEGYRGTAALTVSGKACQAWASQT-----PGDVYS 115
C 94 LGLGKHNYCRNPDNRRPWCYVQVGLKPLVQECMVHDCADG 134
D 116 CQGLVSNYCRNPDGKLPWCYT-----TEYCNVPSCITGG 149

Search completed: May 25, 2004, 14:58:36
Time : 10.2717 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

C protein - protein search, using sw model

R on: May 25, 2004, 14:43:40 ; Search time 5.51446 Seconds
(without alignments)
1350.274 Million cell updates/sec

T le: US-09-880-503-8

S uence: 837
1 SNEHQVNSDCLNGGTCV.....QECMVHPCADKKFSPPEE 143

S ring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

S ched: 141681 seqs, 52070155 residues

T al number of hits satisfying chosen parameters: 141681

M imum DB seq length: 0

M imum DB seq length: 2000000000

F t-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

D abase : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	837	100.0	431	1	UROK_HUMAN
2	757.5	90.5	433	1	UROK_FAPCY
3	625.5	74.7	442	1	UROK_PIG
4	605	72.3	433	1	UROK_BOVIN
5	592	70.7	432	1	UROK_RAT
6	569	68.0	433	1	UROK_MOUSE
7	335.5	40.1	477	1	URT2_DESRO
8	334.5	40.0	434	1	UROK_CHICK
9	328.5	39.2	431	1	URT8_DESRO
10	325	38.8	559	1	TPA_HUMAN
11	325	37.7	559	1	TPA_MOUSE
12	315.5	37.1	477	1	URT1_DESRO
13	310.5	35.9	566	1	TPA_BOVIN
14	300.5	32.0	603	1	FA12_CAVPO
15	268	31.3	655	1	HGFA_HUMAN
16	262	29.9	653	1	HGFA_MOUSE
17	250.5	27.8	615	1	FA12_HUMAN
18	233	26.6	593	1	URT3_DESRO
19	223	25.9	593	1	FA12_BOVIN
20	217	25.2	685	1	ROB1_DROME
21	169	20.2	4548	1	APOA_HUMAN
22	162.5	19.4	473	1	KRM1_RAT
23	161	19.2	1420	1	APOA_MACMU
24	161	19.1	810	1	PLMN_MACMU
25	159.5	19.1	452	1	KRM1_XENLA
26	159	19.0	473	1	KRM1_MOUSE
27	159	19.0	812	1	PLMN_BOVIN
28	159	18.9	937	1	ROR1_HUMAN
29	158.5	18.9	937	1	ROR1_MOUSE
30	158.5	18.9	810	1	PLMN_HUMAN
31	156.5	18.7	475	1	KRM1_HUMAN
32	156	18.6	790	1	PLMN_PIG
33	155	18.5			

ALIGNMENTS

RESULT 1

ID	UROK_HUMAN	STANDARD;	PRT;	431 AA.
AC	P00739; Q15944; Q16618; Q969W6;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	20-MAR-1987 (Rel. 04, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)			
DE	(U-plasminogen activator)			
GN	PLAU.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85215647; PubMed=2987867;			
RA	Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;			
RT	"The human urokinase-plasminogen activator gene and its promoter.";			
RL	Nucleic Acids Res. 13:2759-2771(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,			
RT	Steffens G.J., Heyneker H.L.;			
RA	"Cloning and expression of the gene for pro-urokinase in Escherichia			
RT	coli."			

34	155	18.5	812	1	PLMN_MOUSE
35	153.5	18.3	810	1	PLMN_BRIEU
36	152.5	18.2	169	1	PLMN_RAT
37	150	17.9	711	1	HGFL_HUMAN
38	148.5	17.7	462	1	KRM2_HUMAN
39	148	17.7	333	1	PLMN_CANFA
40	148	17.7	943	1	ROR2_HUMAN
41	148	17.7	944	1	ROR2_MOUSE
42	147.5	17.6	716	1	HGFL_MOUSE
43	146.5	17.5	461	1	KRM2_MOUSE
44	144.5	17.3	728	1	HGF_HUMAN
45	142	17.0	728	1	HGF_MOUSE
46	142	17.0	728	1	HGF_RAT
47	136	16.2	325	1	PLMN_PETMA
48	125.5	15.0	618	1	THRB_MOUSE
49	124.5	14.9	343	1	PLMN_SHEEP
50	124	14.8	622	1	THRB_HUMAN
51	123.5	14.8	724	1	ROR2_DROME
52	117	14.0	2524	1	NOTC_XENLA
53	115	13.7	625	1	THRB_BOVIN
54	113.5	13.6	1238	1	JAG2_HUMAN
55	113	13.3	617	1	THRB_RAT
56	111	13.3	338	1	PLMN_HORSE
57	106.5	12.7	2471	1	NTC3_HUMAN
58	105.5	12.6	1213	1	JAG3_BRARE
59	105.5	12.6	1242	1	JAG1_BRARE
60	105.5	12.6	2109	1	PGCA_CHICK
61	104.5	12.5	2139	1	CRB_DROME
62	103.5	12.4	1321	1	PGCN_HUMAN
63	103	12.3	1295	1	GLP1_CAEEL
64	102	12.3	2470	1	NTC2_MOUSE
65	102	12.2	618	1	DLJ3_HUMAN
66	102	12.2	1202	1	JAG2_RAT
67	102	12.2	1376	1	CRBH_HUMAN
68	102	12.2	2531	1	NTC1_RAT
69	101.5	12.1	686	1	DLJ4_MOUSE
70	101.5	12.1	2556	1	NTC1_HUMAN
71	101.5	12.1	2703	1	NOTC_DROME
72	101	12.1	379	1	WIF1_HUMAN
73	101	12.1	2471	1	NTC2_RAT
74	101	12.1	3014	1	CLR1_HUMAN
75	100	11.9	2003	1	NTC4_HUMAN

P20918	mus musculus
Q29485	erinaceus e
Q01177	rattus norv
Q8ncw0	homo sapien
P80009	canis fami
Q01974	homo sapien
Q92138	mus musculus
P26928	mus musculus
Q8k157	mus musculus
P14210	homo sapien
Q08048	mus musculus
P17945	rattus norv
P33574	petromyzon
P19221	mus musculus
P81286	ovis aries
P00735	bos taurus
P21783	xenopus lae
Q9v8k3	drosophila
Q00734	homo sapien
Q94721	homo sapien
Q90554	brachydanio
Q90577	brachydanio
P07898	gallus gall
P10040	drosophila
O14594	homo sapien
P13508	caenorhabdi
O35516	mus musculus
Q9nyj7	homo sapien
P97607	rattus norv
P82279	homo sapien
Q07008	rattus norv
Q9j171	mus musculus
P46531	homo sapien
P07207	drosophila
Q9y3w5	homo sapien
Q9y3w0	rattus norv
Q9nyv6	homo sapien
Q99466	homo sapien

RL Biotechnology 3:923-929(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86056954; PubMed=2415429;
RA Nagai M., Hiratsuka T., Kaneda T., Hayasuke N., Arimura H.,
RA Nishida M., Suyama T.,
RT "Molecular cloning of cDNA coding for human preprothrombinase";
RL Gene 36:183-188(1985).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=85203359; PubMed=3888571;
RA Jacobs P., Cravador A., Loriau R., Brockly P., Colau B., Chuchana P.,
RA van Elsen A., Herzog A., Bollen A.,
RT "Molecular cloning, sequencing, and expression in Escherichia coli of
RL human prethrombinase cDNA";
RN DNA 4:139-146(1985).
RN [5]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.,
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Peigold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faney J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [7]
RP SEQUENCE OF 66-431 FROM N.A.
RX MEDLINE=84272706; PubMed=6589520;
RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.;
RT "Identification and primary sequence of an unspliced human urokinase
RN poly(A)+ RNA";
RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
RN [8]
RP SEQUENCE OF 21-177.
RX MEDLINE=83055084; PubMed=6754569;
RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
RA Flohe L.,
RT "The primary structure of high molecular mass urokinase from human
RN urine. The complete amino acid sequence of the A chain";
RN Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
RN [9]
RP SEQUENCE OF 156-176 AND 179-224.
RX MEDLINE=83003608; PubMed=6749491;
RA Schaller U., Nick H., Rickli E.E., Gillesse D., Lergier W.,
RA Studer R.O.;
RT "Human low-molecular-weight urinary urokinase. Partial
RN characterization and preliminary sequence data of the two polypeptide
RN chains";
RN Eur. J. Biochem. 125:251-257(1982).
RN [10]
RP SEQUENCE OF 158-410.
RX MEDLINE=83055099; PubMed=6754572;
RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;
RA

RT "The complete amino acid sequence of low molecular mass urokinase
RL from human urine";
RN Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=96000858; PubMed=8591045;
RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
RA Dobson C.M., Stuart D.I., Jones E.Y.;
RT "The crystal structure of the catalytic domain of human
RN urokinase-type plasminogen activator";
RN Structure 3:681-691(1995).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
RX MEDLINE=20266327; PubMed=10805774;
RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
RA Bode W., Magdolen V., Huber R., Moroder L.;
RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
RN selective inhibitors of human urokinase";
RN Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
RN [13]
RP STRUCTURE BY NMR.
RX MEDLINE=89127526; PubMed=2536903;
RA Oswald R.B., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;
RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
RN dimensional NMR";
RL Nature 337:579-582(1989).
RN [14]
RP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=93003110; PubMed=1327118;
RA Li X., Smith R.A.G., Dobson C.M.;
RT "Sequential 1H NMR assignments and secondary structure of the kringles
RN domain from urokinase";
RL Biochemistry 31:9562-9571(1992).
RN [15]
RP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=94149701; PubMed=8107091;
RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
RT "Solution structure of the kringles domain from urokinase-type
RN plasminogen activator";
RL J. Mol. Biol. 235:1548-1559(1994).
RN [16]
RP VARIANT LEU-141.
RX MEDLINE=96186279; PubMed=8652631;
RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
RA Sawasaki Y., Hanada K.;
RT "Characterization of single chain urokinase-type plasminogen
RN activator with a novel amino-acid substitution in the kringles
RL structure";
RN Biochim. Biophys. Acta 1293:83-89(1996).
RN [17]
RP VARIANT LEU-141.
RX MEDLINE=97218551; PubMed=9065988;
RA Conne B., Berczy M., Belin D.;
RT "Detection of polymorphisms in the human urokinase-type plasminogen
RN activator gene";
RL Thromb. Haemost. 77:434-435(1997).
RN [18]
RP ERRATUM.
RA Conne B., Berczy M., Belin D.;
RL Thromb. Haemost. 78:973-973(1997).
RN [19]
RP VARIANT LEU-141.
RX MEDLINE=97337920; PubMed=9194591;
RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,
RA Creutzburg S., Graeff H., Magdolen V.;
RT "Mutational analysis of the genes encoding urokinase-type plasminogen
RN activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer";
RL Electrophoresis 18:686-689(1997).
RN CC -!- FUNCTION: Potent plasminogen activator and is clinically used for
CC therapy of thrombolytic disorders.
RN CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
RN CC -!- SUBUNIT: Found in high and low molecular mass forms. Each consists

of two chains, A and B. The high molecular mass form contains a long chain A. Cleavage occurs after residue 155 in the low molecular mass form to yield a short A1 chain.

-!- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used in Pulmonary Embolism (PE) to initiate fibrinolysis.

-!- SIMILARITY: Belongs to peptidase family S1.

-!- SIMILARITY: Contains 1 EGF-like domain.

-!- SIMILARITY: Contains 1 kringle domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X02419; CAA26368.1; -
EMBL; M15476; AAA61253.1; -
EMBL; D00244; BAA00175.1; -
EMBL; D11143; BAA01919.1; -
EMBL; X02760; CAA26835.1; -
EMBL; AF377330; AAK33822.1; -
EMBL; BC013575; AAH13575.1; -
EMBL; K03226; AAC97138.1; -
EMBL; K02286; AAA61252.1; -
EMBL; A21571; CAA01559.1; -
EMBL; A18397; CAA01390.1; -
PIR; A00931; UKHU
PDB; 1KDU; 31-OCT-93.

Query Match 100.0%; Score 837; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 5e-74;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SNELHGVNSCDLNGGTCVSNKYFSTHWCNCKKFGGQCEIDKSKTCYEGNGHYRG 60
21 SNELHGVNSCDLNGGTCVSNKYFSTHWCNCKKFGGQCEIDKSKTCYEGNGHYRG 80
61 KASTDTMGRCPLPWNASATVLAQTYHAHRSALQGLGKHNYCRNPNRRRWCYVQVGLK 120
81 KASTDTMGRCPLPWNASATVLAQTYHAHRSALQGLGKHNYCRNPNRRRWCYVQVGLK 140
121 PLVQECMVHDCADGKKPSPPPEE 143
141 PLVQECMVHDCADGKKPSPPPEE 163

ILT 2
U: PACY
I: UROK_PAPCY STANDARD; PRT; 433 AA.
P16227;
01-APR-1990 (Rel. 14, Created)
01-APR-1990 (Rel. 14, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
(U-plasminogen activator).
PLAU.
Papio cynocephalus (Yellow baboon).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.
NCBI_TaxID=9556;
[1]_TaxID=9556;
SEQUENCE FROM N.A.
TISSUE:Thoracic aorta;
MEDLINE=90287734; PubMed=2113276;
Au Y.P.T., Wang T.W., Clowes A.W.;
"Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen activator."
Nucleic Acids Res. 18:3411-3411(1990).
-!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

-!- SUBUNIT: Found in high and low molecular mass forms. Each consists of two chains, A and B. The high molecular mass form contains a long chain A. Cleavage occurs after residue 155 in the low molecular mass form to yield a short A1 chain (By similarity).

-!- SIMILARITY: Belongs to peptidase family S1.

-!- SIMILARITY: Contains 1 EGF-like domain.

-!- SIMILARITY: Contains 1 kringle domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X51935; CAA36200.1; -
PIR; S14687; UKBAY.
HSP; P00749; ILWV.
MEROPS; S01.231; -
InterPro; IPR009003; Cys_Ser_trypsin.
InterPro; IPR06209; EGF-like.
InterPro; IPR06210; IEGF.
InterPro; IPR00001; Kringle.
InterPro; IPR008293; Pept_S1A_uPA.
InterPro; IPR001254; Peptidase_S1.
InterPro; IPR001314; Peptidase_S1A.
Pfam; PR00051; kringle; 1.
Pfam; PR00089; trypsin; 1.
PIRSP; PIRSF001144; Urokinase_act; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; KRINGLE.
ProDom; PD000395; Kringle; 1.
SMART; SM00181; EGF; 1.
SMART; SM00130; KR; 1.
SMART; SM00020; Tryp_Spc; 1.
PROSITE; PS00022; EGF_1; 1.
PROSITE; PS01186; EGF_2; FALSE_NEG.
PROSITE; PS50026; EGF_3; 1.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS50070; KRINGLE_2; 1.
PROSITE; PS00240; TRYPsin_DOM; 1.
PROSITE; PS00134; TRYPsin_His; 1.
PROSITE; PS00135; TRYPsin_Ser; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 26 62 EGF-LIKE.
FT DOMAIN 69 150 KRINGLE.
FT DOMAIN 151 177 CONNECTING PEPTIDE.
FT DOMAIN 178 433 SERINE PROTEASE.
FT DISULFID 30 38 BY SIMILARITY.
FT DISULFID 32 50 BY SIMILARITY.
FT DISULFID 52 61 BY SIMILARITY.
FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
FT DISULFID 208 224 BY SIMILARITY.
FT DISULFID 216 287 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
FT CARBOHYD 324 324 N-LINKED (GLCNAC...), (BY SIMILARITY).
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match 90.5%; Score 757.5; DB 1; Length 433;
Best Local Similarity 92.3%; Pred. No. 2.6e-66;
Matches 132; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 SNEHQVPSNCDCLNGTGVSNKYFSNIHWNCNPKKFGQHQCEIDKSKTCYEGNGHGYRG 60
 Db 21 SREL-QVPSDCGCLNGTGVSNKYFSNIHWNCNPKKFGQHQCEIDKSKTCYEGNGHGYRG 79
 QY 61 KASTDTMGRCPLPWNATVLTQYTHAHRSDALQGLGKHNCRNPNRRPWCYVQVGLK 120
 Db 80 KASTDTMGRCPLPWNATVLTQYTHAHRSDALQGLGKHNCRNPNRRPWCYVQVGLK 139
 QY 121 PLVQECWVHDCADGKKPSPPEE 143
 Db 140 QRVOECWVHDCADGKKPSPPEE 162
 RE JT 3
 UR PIG
 ID UROK PIG STANDARD; PRT; 442 AA.
 AC P04185;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator)
 GN PLAU.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=85087954; PubMed=6096832;
 RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
 RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
 RL Nucleic Acids Res. 12:9525-9541(1984).
 RN [2]
 RP REVISION TO 241.
 RA Nagamine Y.;
 RL Submitted (DEC-1986) to the PIR data bank.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X01648; CAA25806.1; -;
 CC EMBL; X02724; CAA26511.1; -;
 CC PIR; A00932; UKPG.
 CC HSP; P00749; iKDU.
 CC MROPS; S01.233; -;
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR006209; EGF-like.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR008293; Pept_S1A_uPA.
 CC InterPro; IPR001254; Peptidase_S1.
 CC InterPro; IPR001314; Peptidase_S1A.
 CC Pfam; PF00051; kringle; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PIRSF; PIRSF001144; Utk_plasm_act; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; Kringle; 1.
 CC SMART; SM00130; KR; 1 SPC.
 CC SMART; SM00020; TrypSPC; 1.
 CC PROSITE; PS00022; EGF_1; 1.

DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PROSITE; PS00026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00133; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
 FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
 FT DOMAIN 29 65 EGF-LIKE.
 FT DOMAIN 72 153 KRINGLE.
 FT DOMAIN 154 189 CONNECTING PEPTIDE.
 FT DOMAIN 190 442 SERINE PROTEASE.
 FT CARBOHYD 152 152 N-LINKED (GLCNAC...).
 FT DISULFID 33 41 BY SIMILARITY.
 FT DISULFID 35 53 BY SIMILARITY.
 FT DISULFID 55 64 BY SIMILARITY.
 FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 220 236 BY SIMILARITY.
 FT DISULFID 228 299 BY SIMILARITY.
 FT DISULFID 324 393 BY SIMILARITY.
 FT DISULFID 356 372 BY SIMILARITY.
 FT DISULFID 383 411 BY SIMILARITY.
 FT ACT_SITE 235 CHARGE RELAY SYSTEM.
 FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
 FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
 FT CONFLICT 241 241 Q -> H (IN REF. 1; CAA25806).
 FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA26511).
 FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
 SQ SEQUENCE 442 AA; 49116 MW; E32FCEP501321EE CRC64;
 Query Match 74.7%; Score 625.5; DB 1; Length 442;
 Best Local Similarity 73.4%; Pred. No. 1.8e-53;
 Matches 113; Conservative 11; Mismatches 19; Indels 11; Gaps 2;
 QY 1 SNEHQV--PSNCDCLNGTGVSNKYFSNIHWNCNPKKFGQHQCEIDKSKTCYEGNGHGY 58
 Db 21 SHELHQESGASNGCLNGKGVSKYFSNIQRSCPKKFGQHQCEIDTSQTCFEGNGHSY 80
 QY 59 RGKASTDTMGRCPLPWNATVLTQYTHAHRSDALQGLGKHNCRNPNRRPWCYVQV 118
 Db 81 RGKANTNTGRCPLPWNATVLTQYTHAHRSDALQGLGKHNCRNPNRRPWCYVQV 140
 QY 119 LKPLVQECWVHDCA-----DGKKPSPPEE 143
 Db 141 LKQLVQECWVHDCGSHRPAIDYDGNPFTPEK 174
 RESULT 4
 ID UROK BOVIN STANDARD; PRT; 433 AA.
 AC Q05589; Q05209;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator)
 GN PLAU.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aortic endothelium;
 RX MEDLINE=93216119; PubMed=8385052;
 RA Kretschmar J., Haendler B., Kojima S., Rifkin D.B.,
 RA Schlemming W.-D.;

FT ACT SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ CONFLICT 189 A -> T (IN REF. 2).
 SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 72.3%; Score 605; DB 1; Length 433;
 Best Local Similarity 72.4%; Pred. No. 1.7e-51;
 Matches 105; Conservative 14; Mismatches 24; Indels 2; Gaps 1;

QY 1 SNEHQV--PSNCDCLNGTCTVSNKYFNSNIHWCNCPKFGQGHCEIDSKTCYEGNGHFF 58
 DB 21 SNEVHKGESGCGCLNGKCKVYKFSNIQRCSPKFGQGHCEIDSKTCYQNGHSHY 80
 QY 59 RGAATDTMGRCLPWNSTVLTQYHAHRSDALQGLGKHNKYNCRPNRRPWCYVQVG 118
 DB 81 RGNARDLGSRFLAWDSPTVLLKMYHAHRSDAIQGLGKHNKYNCRPNRRPWCYVQVG 140
 QY 119 LKPLVQECMVHDCADGKPPSPPEE 143
 DB 141 LKQVQFCVQDCSVGKSPSPPREK 165

RESULT 5
 UROK RAT STANDARD; PRT; 432 AA.

ID UROK RAT STANDARD; PRT; 432 AA.

AC P29598;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Fischer 344;
 RX MEDLINE=92233409; PubMed=1568219;
 RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
 RA Kefford R.F.;
 RT "Transcriptional and posttranscriptional activation of urokinase
 RT plasminogen activator gene expression in metastatic tumor cells";
 RL Cancer Res. 52:2489-2496(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Rabbani S.A.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -|- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 156 in the low
 CC molecular mass form to yield a short A1 chain (By similarity).
 CC -|- SIMILARITY: Belongs to peptidase family S1.
 CC -|- SIMILARITY: Contains 1 EGF-like domain.
 CC -|- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X63434; CAA45028.1; -;
 DR EMBL; X65651; CAA45601.1; -;
 DR PIR; S24604; S18932.
 DR HSSP; P00749; IKDU.
 DR MEROPS; S01.231; -;
 DR InterPro; IPR009003; Cys_Ser_trypsin.

"Bovine urokinase-type plasminogen activator and its receptor:
 cloning and induction by retinoic acid.";
 Gene 125:177-183(1993).
 [2]

SEQUENCE OF 12-433 FROM N.A.
 TISSUE=Kidney;
 Ravn P., Berglund L., Petersen T.B.;
 "Cloning and characterization of the bovine plasminogen activators uPA
 and tPA";
 Int. Dairy J. 5:605-617(1995).
 C -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 C plasminogen to form plasmin.
 C -|- INDUCTION: By retinoic acid.
 C -|- SIMILARITY: Belongs to peptidase family S1.
 C -|- SIMILARITY: Contains 1 EGF-like domain.
 C -|- SIMILARITY: Contains 1 kringle domain.
 C -----
 C This SWISS-PROT entry is copyright. It is produced through a collaboration
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -
 C the European Bioinformatics Institute. There are no restrictions on its
 C use by non-profit institutions as long as its content is in no way
 C modified and this statement is not removed. Usage by and for commercial
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 C or send an email to license@isb-sib.ch).
 C -----
 C EMBL; L03546; AAA51419.1; -;
 DR EMBL; X85801; CAA59796.1; -;
 DR PIR; J00560; J00560.
 DR HSSP; P00749; ILMW.
 DR MEROPS; S01.231; -;
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR008293; Peptidase_S1A.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PIRSF001144; Urk_plasm act; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; Kringle; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; KG; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PROSITE; PS00026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 DR Kringle; EGF-like domain; Signal; Zymogen.
 DR SIGNAL; 1 20
 F CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 F CHAIN 21 179 CHAIN A (BY SIMILARITY).
 F CHAIN 181 433 CHAIN B (BY SIMILARITY).
 F DOMAIN 29 65 EGF-LIKE.
 F DOMAIN 72 153 KRINGLE.
 F DOMAIN 154 180 CONNECTING PEPTIDE.
 F DOMAIN 181 433 SERINE PROTEASE.
 F DISULFID 33 41 BY SIMILARITY.
 F DISULFID 35 53 BY SIMILARITY.
 F DISULFID 55 64 BY SIMILARITY.
 F DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
 F DISULFID 211 227 BY SIMILARITY.
 F DISULFID 219 290 BY SIMILARITY.
 F DISULFID 315 384 BY SIMILARITY.
 F DISULFID 347 363 BY SIMILARITY.
 F DISULFID 374 402 BY SIMILARITY.
 F ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
 F ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).

F SIGNAL 1 20
 F CHAIN 21 433
 F CHAIN 21 178
 F CHAIN 157 178
 F CHAIN 180 433
 F CHAIN 28 64
 F DOMAIN 71 152
 F DOMAIN 153 179
 F DOMAIN 180 433
 F DISULFID 32 40
 F DISULFID 34 52
 F DISULFID 54 63
 F DISULFID 169 301
 F DISULFID 211 227
 F DISULFID 219 290
 F DISULFID 315 384
 F DISULFID 347 363
 F DISULFID 374 402
 F ACT SITE 226 226
 F ACT SITE 277 277
 F ACT SITE 378 378
 S SEQUENCE 433 AA; 48268 MW; A99C35F6250443F9 CRC64;

very Match 68.0%; Score 569; DB 1; Length 433;
 est Local Similarity 70.4%; Pred No. 5.3e-48;
 ices 95; Conservative 15; Mismatches 25; Indels 0; Gaps 0;

Q 9 SNCDCLNGTGVSNKYFNIHWCNPKFGQGHCEIDSKTCYEGNGHFGYRGKASTDTMG 68
 D 30 SNCQCGQGVCSYKYFGRIRRCSPKRFQGEHCEIDASKTCYHNGSDSYRGKANTDTKG 89
 Q 69 RPLCPWNSATVLQOYTHAHRSDALQLGLGKHNCRNPNRNRPRPCYQVGLKPLAVQSCMV 128
 D 90 RPLCAWNAVLQKPYNAHRPDAISLGLGKHNCRNPNQKRPWCYQVGLRQVQSCMV 149
 Q 129 HDCADGKFPSPPEE 143
 D 150 HDCSLSKXFPSSVDQ 164

R IT 7
 U DESRO STANDARD; PRT; 477 AA.
 I URT2_DESRO
 A P15638;
 D 01-APR-1990 (Rel. 14, Created)
 D 01-FEB-1996 (Rel. 33, Last sequence update)
 D 28-FEB-2003 (Rel. 41, Last annotation update)
 D Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA
 alpha-2) (BART-PA) (T-plasminogen activator).
 C Desmodus rotundus (Vampire bat).
 O Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 O Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 O Desmodontinae; Desmodus.
 O NCBI_TaxID=9430;
 R [1]_TaxID=9430;
 R SEQUENCE FROM N.A.
 R TISSUE=Salivary gland;
 R MEDLINE=9203036; PubMed=1937019;
 R Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 R Alagon A., Donner P., Schleuning W.D.;
 R "The plasminogen activator family from the salivary gland of the
 R vampire bat Desmodus rotundus: cloning and expression.";
 R Gene 105:229-237 (1991).
 R [2].
 R SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 R TISSUE=Salivary gland;
 R MEDLINE=9003667; PubMed=2509450;
 R Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
 R Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
 R "Isolation, characterization, and cDNA cloning of a vampire bat
 R salivary plasminogen activator.";
 R J. Biol. Chem. 264:17947-17952 (1989).
 R [3]

RP CHARACTERIZATION.
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schleuning W.D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baidus B., Witt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403 (1992).
 CC -I- FUNCTION: Probably essential to support the feeding habits of this
 CC exclusively haematophagous animal. Probable potent thrombolytic
 CC agent.
 CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -I- ENZYME REGULATION: Activity toward plasminogen is stimulated in
 CC the presence of fibrin I.
 CC -I- SUBUNIT: Monomer.
 CC -I- DOMAIN: The fibronectin type-I domain mediates binding to fibrin,
 CC and the kringle domain apparently mediates fibrin-induced
 CC stimulation of activity.
 CC -I- SIMILARITY: Belongs to peptidase family S1.
 CC -I- SIMILARITY: Contains 1 EGF-like domain.
 CC -I- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -I- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M63988; AAA31593.1; -;
 DR EMBL; J05082; AAA31596.1; -;
 DR PIR; A34369; A34369.
 DR PIR; J05098; J05098.
 DR HSSP; P98119; 1A51.
 DR MEROPS; S01.232; -;
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR000083; Fibrnctn1.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00039; fni; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00058; FNI; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPsin DOM; 1.
 DR PROSITE; PS00134; TRYPsin HIS; 1.
 DR PROSITE; PS00135; TRYPsin SER; 1.
 DR PROSITE; PS00135; TRYPsin SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 FT SIGNAL 1 36
 FT CHAIN 37 477
 FT DOMAIN 40 82
 FT DOMAIN 83 121
 FT DOMAIN 128 209
 FT DOMAIN 225 477
 FT SERINE PROTEASE.

FT ACT_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 321 321 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 428 428 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT DISULFID 42 72 BY SIMILARITY.

FT DISULFID 70 79 BY SIMILARITY.

FT DISULFID 87 98 BY SIMILARITY.

FT DISULFID 92 109 BY SIMILARITY.

FT DISULFID 111 120 BY SIMILARITY.

FT DISULFID 128 209 BY SIMILARITY.

FT DISULFID 149 191 BY SIMILARITY.

FT DISULFID 180 204 BY SIMILARITY.

FT DISULFID 214 345 BY SIMILARITY.

FT DISULFID 257 273 BY SIMILARITY.

FT DISULFID 265 334 BY SIMILARITY.

FT DISULFID 359 434 BY SIMILARITY.

FT DISULFID 391 407 BY SIMILARITY.

FT DISULFID 424 452 BY SIMILARITY.

FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 403 403 N -> K (IN REF. 2).

FT CONFLICT 417 417 Y -> H (IN REF. 2).

FT CONFLICT 435 435 M -> R (IN REF. 2).

FT SEQUENCE 477 AA; 53719 MW; 17486555C0E5077C CRC64;

Query Match 40.1%; Score 335.5; DB 1; Length 477;

Best Local Similarity 46.3%; Pred. No. 2.7e-25;

Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;

QY 3 ELHQVP----SNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFY 58

DB 78 QCHTVPKSCSELRCFNGGTCWQAASFDF-VQCPKGYTGQCEVDTHATCYKQGVY 136

QY 59 RGKASTDTMGRPCLPNWSATVLCQTHAHRSDALQLGLGKHYCRNPNRRPWCYVQV 118

DB 137 RGWSTSESQAQINNSNLLTRTYNGRRSDAITLGLGNHYCRNPNNSXPWCYVKA 196

QY 119 LKPLVQECMVHCA 132

DB 197 SKFLEFCSPVCS 210

RE T 8

UR CHICK STANDARD; PRT; 434 AA.

ID P15120; UROK CHICK STANDARD; PRT; 434 AA.

DI 01-APR-1990 (Rel. 14, Created)

DI 01-APR-1990 (Rel. 14, Last sequence update)

DI 28-FEB-2003 (Rel. 41, Last annotation update)

DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)

DE (U-plasminogen activator).

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90110185; PubMed=2295632;

RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;

RT "The chicken urokinase-type plasminogen activator gene.";

RL J. Biol. Chem. 265:1339-1344 (1990).

CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

CC -1- SIMILARITY: Belongs to peptidase family S1.

CC -1- SIMILARITY: Contains 1 EGF-like domain.

CC -1- SIMILARITY: Contains 1 kringle domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

CC EMBL; J05187; AAA49131.1; -.

CC EMBL; J05188; AAA49130.1; -.

DR PIR; A35005; A35005.

DR HSP; P00763; IDPO.

DR MEROPS; S01.231; -.

DR InterPro; IPR009003; Cys Ser trypsin.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR006210; EGF.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR008293; Pept_S1A_uPA.

DR InterPro; IPR01254; Peptidase_S1.

DR InterPro; IPR001314; Peptidase_S1A.

DR Pfam; PF00051; kringle; 1.

DR Pfam; PF00089; trypsin; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PRS001144; Urokinase; 1.

DR PRINTS; PRS00722; CHYMOTRYPSIN.

DR PRINTS; PRS00018; KRINGLE.

DR PRODOM; PD000395; Kringle; 1.

DR SMART; SM00181; EGF; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp_SPC; 1.

DR PROSITE; PS00022; EGF_1; 1.

DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS00026; EGF_3; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

DR PROSITE; PS02040; TRYPSIN_DOM; 1.

DR PROSITE; PS01134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

DR Plasminogen activation; Hydrolase; Serine protease; Glycoprotein; Kringle; EGF-like domain; Signal; Zymogen.

FT SIGNAL 1 20 POTENTIAL.

FT CHAIN 21 434 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.

FT CHAIN 21 171 CHAIN A (BY SIMILARITY).

FT CHAIN 173 434 CHAIN B (BY SIMILARITY).

FT DOMAIN 36 72 EGF-LIKE.

FT DOMAIN 79 158 KRINGLE.

FT DOMAIN 159 172 CONNECTING PEPTIDE.

FT DOMAIN 173 434 SERINE PROTEASE.

FT DISULFID 40 48 BY SIMILARITY.

FT DISULFID 42 60 BY SIMILARITY.

FT DISULFID 62 71 BY SIMILARITY.

FT DISULFID 162 296 INTERCHAIN (BY SIMILARITY).

FT DISULFID 202 218 BY SIMILARITY.

FT DISULFID 210 285 BY SIMILARITY.

FT DISULFID 310 379 BY SIMILARITY.

FT DISULFID 342 358 BY SIMILARITY.

FT DISULFID 369 397 BY SIMILARITY.

FT ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 373 373 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 434 AA; 49400 MW; ED881048DD666A55 CRC64;

Query Match 40.0%; Score 334.5; DB 1; Length 434;

Best Local Similarity 54.2%; Pred. No. 3.1e-25;

Matches 58; Conservative 14; Mismatches 30; Indels 5; Gaps 2;

QY 11 CDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRGKASTDTMGRP 70

DB 40 CQCLNGGTCVSYRFFSQIKRCLCPGEGYGLHCHIDTNSICYSNGNGBDYRMAEDP----G 95

QY 71 CLPWSATVLC-QTYHAHRSDALQLGLGKHYCRNPNRRPWCYVQ 116

DB 96 CLYWDHFSVIRWDYHADLKNALQLGLGKHYCRNPNRSEPRWCYTK 142

RESULT 9

URTB DESRO

ID URTB DESRO STANDARD; PRT; 431 AA.

AC P98121;

I 01-FEB-1996 (Rel. 33, Created)
D 01-FEB-1996 (Rel. 33, Last sequence update)
D 28-FEB-2003 (Rel. 41, Last annotation update)
D Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA
D beta).
C Desmodus rotundus (Vampire bat).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
C Desmodontinae; Desmodus.
C NCBI_TaxID=9430;
R [1]_SEQUENCE FROM N.A.
R TISSUE=Salivary gland;
R MEDLINE=9203036; PubMed=1937019;
R Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
R Alagon A., Donner P., Schleuning W.D.;
R "The plasminogen activator family from the salivary gland of the
R vampire bat Desmodus rotundus: cloning and expression.";
R Gene 105:229-237(1991).
R [2]
R CHARACTERIZATION.
R MEDLINE=9333059; PubMed=1309059;
R Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
R Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
R Donner P.;
R "Plasminogen activators from the saliva of Desmodus rotundus (common
R vampire bat): unique fibrin specificity.";
R Ann. N.Y. Acad. Sci. 667:395-403(1992).
C -!- FUNCTION: Probably essential to support the feeding habits of this
C exclusively haematophagous animal. Probable potent thrombolytic
C agent.
C -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
C plasminogen to form plasmin.
C -!- SUBUNIT: Monomer.
C -!- SIMILARITY: Belongs to peptidase family S1.
C -!- SIMILARITY: Contains 1 EGF-like domain.
C -!- SIMILARITY: Contains 1 kringle domain.
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation.
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
C or send an email to license@isb-sib.ch).
C -----
C EMBL; M63989; AAA31594.1; -
C F01; J80599; J80599.
C HSP; P98119; IAS1.
C MEROPS; S01.239; -.
D InterPro; IPR009003; Cys Ser trypsin.
D InterPro; IPR006209; EGF-like.
D InterPro; IPR006210; IEGF.
D InterPro; IPR000001; Kringle.
D InterPro; IPR001254; Peptidase_S1.
D InterPro; IPR001314; Peptidase_S1A.
D Pfam; PF00008; EGF; 1.
D Pfam; PF00051; kringle; 1.
D Pfam; PF00089; trypsin; 1.
D PRINTS; PR00722; CHYMOTRYPSIN.
D PRINTS; PR00018; KRINGLE.
D ProDom; PD000395; Kringle; 1.
D SMART; SM00181; EGF; 1.
D SMART; SM00130; KS; 1.
D SMART; SM00020; Tryp_Spc; 1.
D PROSITE; PS00022; EGF_1; 1.
D PROSITE; PS01186; EGF_2; 1.
D PROSITE; PS00026; EGF_3; 1.
D PROSITE; PS00021; KRINGLE_1; 1.
D PROSITE; PS00070; KRINGLE_2; 1.
D PROSITE; PS00240; TRYPSIN_DOM; 1.
D PROSITE; PS00134; TRYPSIN_HIS; 1.
D PROSITE; PS00135; TRYPSIN_SER; 1.

KW plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36 POTENTIAL.
FT CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.
FT DOMAIN 37 75 EGF-LIKE.
FT DOMAIN 82 163 KRINGLE.
FT DOMAIN 179 431 SERINE PROTEASE.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 41 52 BY SIMILARITY.
FT DISULFID 46 53 BY SIMILARITY.
FT DISULFID 65 74 BY SIMILARITY.
FT DISULFID 82 163 BY SIMILARITY.
FT DISULFID 103 145 BY SIMILARITY.
FT DISULFID 134 158 BY SIMILARITY.
FT DISULFID 168 299 BY SIMILARITY.
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 288 BY SIMILARITY.
FT DISULFID 313 388 BY SIMILARITY.
FT DISULFID 345 361 BY SIMILARITY.
FT DISULFID 378 406 BY SIMILARITY.
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 431 AA; 48221 MW; 699B5E675B162CBF CRC64;
Query Match 39.2%; Score 328.5; DB 1; Length 431;
Best Local Similarity 47.6%; Pred. No. 1.2e-24;
Matches 59; Conservative 16; Mismatches 48; Indels 1; Gaps 1;
QY 9 SNCDCLNGGTCVSNKYFSNIHWCNCPKFGGHCIDKSKTCVGGNGHYRGKASDTWTG 68
DB 42 SELRCFNGGTCAASFSDF-VCCQPKGYTGQCEVDTHATCYKQGVYRGWTSSEG 100
QY 69 RPLCPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVGLKPLVOECMV 128
DB 101 AQCIWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPDPNNKPCWYKASKFILEFCV 160
QY 129 HDCA 132
DB 161 PVCS 164
RESULT 10
TPA_HUMAN
ID TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103.
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator) (Alteplase) (Reteplase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Penica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A., H.L.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli.";
RL Nature 301:214-221(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=88262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA

from human fetal lung cells.
 Nucleic Acids Res. 15:5695-5695(1988).
 [3]
 SEQUENCE FROM N.A.
 MEDLINE=88054470; PubMed=2824147;
 Reddy V.B., Garramone A.J., Sasek H., Wei C.-M., Watkins P., Galli J.,
 Hsiung N.;
 "Expression of human uterine tissue-type plasminogen activator in
 mouse cells using BPV vectors."
 DNA 6:451-472(1987).
 [4]
 SEQUENCE FROM N.A.
 MEDLINE=86196143; PubMed=3009482;
 Friezen Degen S.J., Rajput B., Reich E.;
 "The human tissue plasminogen activator gene."
 J. Biol. Chem. 261:6972-6985(1986).
 [5]
 SEQUENCE FROM N.A.
 MEDLINE=84298137; PubMed=6089198;
 Ny T., Elgh F., Lund B.;
 "The structure of the human tissue-type plasminogen activator gene:
 correlation of intron and exon structures to functional and
 structural domains."
 Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
 [6]
 SEQUENCE FROM N.A.
 MEDLINE=86284200; PubMed=3030401;
 Harris T.J., Patel T., Marston F.A., Little S., Ertage J.S.,
 Opendakker G., Volckaert G., Rembouts W., Billiau A., Somer P.;
 "Cloning of cDNA coding for human tissue-type plasminogen activator
 and its expression in *Escherichia coli*."
 Mol. Biol. Med. 3:279-292(1986).
 [7]
 SEQUENCE FROM N.A. (ISOFORM SHORT).
 TISSUE=Umbilical vein;
 MEDLINE=90192129; PubMed=2107528;
 Siebert P.D., Fong K.;
 "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
 human endothelial cells."
 Nucleic Acids Res. 18:1086-1086(1990).
 [8]
 SEQUENCE FROM N.A.
 TISSUE=Brain;
 MEDLINE=22388237; PubMed=12477932;
 Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan R., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences."
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [9]
 SEQUENCE OF 212-361 FROM N.A.
 MEDLINE=83169656; PubMed=6572897;
 Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
 Josephson S.;
 "Isolation of cDNA sequences coding for a part of human tissue
 plasminogen activator."
 Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
 [10]

RP SEQUENCE OF 1-36 FROM N.A.
 RX MEDLINE=85289338; PubMed=3161893;
 Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
 Schleuning W.-D.;
 "Isolation and characterization of the human tissue-type plasminogen
 activator structural gene including its 5' flanking region."
 J. Biol. Chem. 260:11223-11230(1985).
 [11]
 RP SEQUENCE OF 31-562 FROM N.A.
 RX MEDLINE=91291340; PubMed=1368681;
 Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
 "Purification and characterization of tissue plasminogen activator
 secreted by human embryonic lung diploid fibroblasts, IMR-90 cells."
 Agric. Biol. Chem. 55:1225-1232(1991).
 [12]
 RN "Establish glycosylation sites, and localize variant differences."
 Biochemistry 23:3701-3707(1984).
 [13]
 RP SEQUENCE OF 36-562.
 TISSUE=Melanoma;
 RX MEDLINE=85000468; PubMed=6433976;
 Pohl G., Kaellstroem M., Bergsdorf N., Wallen P., Joernvall H.;
 "Tissue plasminogen activator: peptide analyses confirm an indirectly
 derived amino acid sequence, identify the active site serine residue,
 establish glycosylation sites, and localize variant differences."
 Biochemistry 23:3701-3707(1984).
 [14]
 RP SEQUENCE OF 33-52 AND 311-330.
 TISSUE=Melanoma;
 RX MEDLINE=83209620; PubMed=6682760;
 Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
 "Purification and characterization of a melanoma cell plasminogen
 activator."
 Eur. J. Biochem. 132:681-686(1983).
 [15]
 RP CARBOHYDRATE-LINKAGE SITE THR-96.
 MEDLINE=91159408; PubMed=1900431;
 Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
 "Tissue plasminogen activator has an O-linked fucose attached to
 threonine-61 in the epidermal growth factor domain."
 Biochemistry 30:2311-2314(1991).
 [16]
 RP DISULFIDE BONDS IN KRINGLE 2.
 MEDLINE=91244765; PubMed=1645336;
 Vlahos C.J., Wilhelm O.G., Hassell T., Jaekunas S.R., Bang N.U.;
 "Disulfide pairing of the recombinant kringle-2 domain of tissue
 plasminogen activator produced in *Escherichia coli*."
 J. Biol. Chem. 266:10070-10072(1991).
 [17]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
 MEDLINE=95200985; PubMed=9613982;
 Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
 Bode W.;
 "The 2.3 Å crystal structure of the catalytic domain of recombinant
 two-chain human tissue-type plasminogen activator."
 J. Mol. Biol. 259:117-135(1996).
 [18]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
 MEDLINE=97449126; PubMed=9305622;
 Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
 Bode W.;
 "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
 crystal structure of single-chain human tPA."
 EMBO J. 16:4797-4805(1997).
 [19]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
 MEDLINE=92118803; PubMed=1310033;
 de Vos A., Ultsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
 Westbrook M.L., Kossiakof A.A.;

Pfam; PF00089; trypsin; 1.
 PRINTS; PRO0722; CHYMOTRYPSIN.
 PRINTS; PRO0018; KRINGLE.
 ProDom; PD000395; Kringle; 2.
 SMART; SM00181; EGF; 1.
 SMART; SM00058; FNI; 1.
 SMART; SM00130; KR; 2.
 SMART; SM00020; TRYP SPG; 1.
 PROSITE; PS00022; EGF 1; 1.
 PROSITE; PS01186; EGF 2; 1.
 PROSITE; PS01253; FIBRONECTIN 1; 1.
 PROSITE; PS00026; EGF 3; 1.
 PROSITE; PS00021; KRINGLE 2; 2.
 PROSITE; PS00070; KRINGLE 2; 2.
 PROSITE; PS00240; TRYP SIN DOM; 1.
 PROSITE; PS00134; TRYP SIN HIS; 1.
 PROSITE; PS00135; TRYP SIN SER; 1.
 Plasmogen activation; Hydrolase; Serine protease; Glycoprotein;
 Plasma; Kringle; EGF-like domain; Repeat; Signal.
 SIGNAL 1 17 PROBABLE.
 PROPEP 18 29
 CHAIN 30 559
 CHAIN 30 308
 CHAIN 309 559
 DOMAIN 36 78
 DOMAIN 79 107
 DOMAIN 124 205
 DOMAIN 213 294
 DOMAIN 309 559
 ACT SITE 355 355
 ACT SITE 404 404
 ACT SITE 510 510
 DISULFID 38 68
 DISULFID 66 75
 DISULFID 83 94
 DISULFID 88 105
 DISULFID 107 116
 DISULFID 124 205
 DISULFID 145 187
 DISULFID 176 200
 DISULFID 213 294
 DISULFID 234 276
 DISULFID 265 289
 DISULFID 297 428
 DISULFID 340 356
 DISULFID 348 417
 DISULFID 442 516
 DISULFID 474 490
 DISULFID 506 534
 CARBOHYD 149 149
 CARBOHYD 481 481
 CONFLICT 260 260
 CONFLICT 325 325
 SEQUENCE 559 AA; 63122 MW; 8CEE2BDB94514D9 CRC64;
 Query Match 37.7%; Score 315.5; DB 1; Length 559;
 est Local Similarity 44.5%; Pred. No. 2.8e-23;
 atches 61; Conservative 15; Mismatches 56; Indels 5; Gaps 2;
 3 ELHQVP-----SNCDCLNGCTVSNKYFNIHWNCNPKFGGCHCIDSIDKTCYEGNGHFY 58
 74 QCHSVPRVRCSPRCFNGGTCCQALYFSDP-VCCPDPGVGKRCIDIDTRATCFEQGITY 132
 59 RGKASTDTMGRCPLPWNSTVLTQTYVHRSVALQLGLGKHNCRPNRRPWCYQVVG 118
 133 RGWTSTAESGAECINWSSVLSLKPYNARRPNAIKLGLGNHNYCRNPRDLKPCYVFK 192
 119 LKPLVQECMWHCADGK 135
 193 GKYTFECSTPACPKG 209

RESULT 13
 URT1_DESRO STANDARD; PRT; 477 AA.
 AC P98119; 1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSPA
 alpha-1).
 OS Desmodus rotundus (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomae;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 OX NCBI_TaxID=9430;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=9339059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
 Donner P.;
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RC TISSUE=Salivary gland;
 RX MEDLINE=9802741; PubMed=9354616;
 RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,
 Schleuning W.D., Bode W.;
 RT "Catalytic domain structure of vampire bat plasminogen activator: a
 RT molecular paradigm for proteolysis without activation cleavage.";
 RL Biochemistry 36:13483-13493(1997).
 CC -!- FUNCTION: Probably essential to support the feeding habits of this
 CC exclusively haematophagous animal. Potent thrombolytic agent.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- ENZYME REGULATION: Activity toward plasminogen is stimulated in
 CC the presence of fibrin I.
 CC -!- SUBUNIT: Monomer.
 CC -!- DOMAIN: The fibronectin type-I domain mediates binding to fibrin,
 CC and the kringle domain apparently mediates fibrin-induced
 CC stimulation of activity.
 CC -!- SIMILARITY: Belongs to peptidase family S1.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M63987; AAA31591.1; -;
 CC EMBL; M63986; AAA31592.1; -;
 CC PIR; J05097; J05097.
 CC PDB; 1A51; 23-MAR-99.
 CC MEROPS; S01.232; -;
 CC GlycoSuiteDB; P98119; -;
 CC InterPro; IPR009003; Cys Ser trypsin.
 CC InterPro; IPR006209; EGF_like.

DR InterPro: IPR000083; Fibrinctnl.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00039; fnl; 1.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00058; fnl; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00026; EGF_3; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.
FT SIGNAL 1 36
FT CHAIN 37 477
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 225 477
FT ACT_SITE 272 272
FT ACT_SITE 321 321
FT ACT_SITE 428 428
FT DISULFID 42 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 214 245
FT DISULFID 257 273
FT DISULFID 265 334
FT DISULFID 359 434
FT DISULFID 391 407
FT DISULFID 424 452
FT CARBOHYD 153 153
FT CARBOHYD 398 398
FT TURN 214 215
FT STRAND 223 224
FT TURN 226 227
FT STRAND 230 231
FT HELIX 234 236
FT TURN 238 239
FT STRAND 240 245
FT STRAND 254 263
FT TURN 264 265
FT STRAND 266 269
FT HELIX 271 273
FT TURN 280 282
FT STRAND 284 287
FT TURN 297 298
FT STRAND 300 309
FT TURN 311 312
FT TURN 315 317
FT TURN 319 320
FT STRAND 323 328

FT STRAND 338 338
FT TURN 339 340
FT STRAND 341 341
FT STRAND 345 345
FT TURN 349 350
FT TURN 355 356
FT STRAND 358 363
FT STRAND 366 366
FT STRAND 374 374
FT STRAND 379 385
FT HELIX 388 390
FT TURN 393 398
FT TURN 403 404
FT STRAND 405 409
FT TURN 425 426
FT TURN 428 429
FT STRAND 431 436
FT TURN 437 438
FT STRAND 439 448
FT TURN 455 456
FT STRAND 459 463
FT HELIX 464 467
FT HELIX 468 474
SQ SEQUENCE 477 AA; 53616 MW; AA06FD1739C10B5E CRC64;
Query Match 37.1%; Score 310.5; DB 1; Length 477;
Best Local Similarity 45.5%; Pred. No. 7.2e-23;
Matches 60; Conservative 14; Mismatches 53; Indels 5; Gaps 3;
QY 5 HQVPSN-CD---CLNGGTCVSNKYFSNIHWCNCPKKGGOHCIDKSKTCYEGNGHYRG 60
Db 80 HTVPVNSCEPRCNGGTCQAVYFSDF-VCCPAGTGRKCEVDTRATCYEGGVTYRG 138
QY 61 KASDTTMRPCLPWN SATVLQOTYHAHRS DALQGLGKHNCRNPONRRFPWCYVQVGLK 120
Db 139 TWSTAESRVECI NNWSLLTTRTYNGRMPDAFNGLGNHNYCRNPNGAPKPCVVIKAGK 198
QY 121 PLVQECMVHDC A 132
Db 199 FTSESCVPVCS 210
RESULT 14
TPA_BOVIN STANDARD; PRT; 566 AA.
AC Q26198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
and tPA";
RL Int. Dairy J. 5:605-617(1995).
CC -!- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen
to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By
controlling plasmin-mediated proteolysis, it plays an important
role in tissue remodeling and degradation, in cell migration and
many other physiological events.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -!- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide
bond.

-1- SUBCELLULAR LOCATION: Secreted; extracellular.
 -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
 -1- MISCELLANEOUS: Binds to the kringle structure of the fibrin A chain. Binding to fibrin enhances its catalytic activity.
 -1- SIMILARITY: Belongs to peptidase family S1.
 -1- SIMILARITY: Contains 1 EGF-like domain.
 -1- SIMILARITY: Contains 1 fibronectin type I domain.
 -1- SIMILARITY: Contains 2 kringle domains.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X85800; CAA59795.1; ..
 HSSP; P00750; IRTF.
 MEROPS; S01.234; ..
 InterPro; IPR009003; Cys_Ser_trypsin.
 InterPro; IPR006209; EGF-like.
 InterPro; IPR000083; Fibnctn1.
 InterPro; IPR006210; IEGF.
 InterPro; IPR000001; Kringle.
 InterPro; IPR001254; Peptidase_S1.
 InterPro; IPR001314; Peptidase_S1A.
 Pfam; PF00008; EGF; 1.
 Pfam; PF00039; fni; 1.
 Pfam; PF00051; kringle; 2.
 Pfam; PF00089; trypsin; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 ProDom; PD000395; Kringle; 2.
 SMART; SM00181; EGF; 1.
 SMART; SM00058; FN1; 1.
 SMART; SM00130; KR; 2.
 SMART; SM00020; TYP_SPC; 1.
 PROSITE; PS00022; EGF_1; 1.
 PROSITE; PS01186; EGF_2; 1.
 PROSITE; PS00026; EGF_3; 1.
 PROSITE; PS01253; FIBRONECTIN_1; 1.
 PROSITE; PS00021; KRINGLE_1; 1.
 PROSITE; PS00070; KRINGLE_2; 2.
 PROSITE; PS50240; TRYPSIN_DOM; 1.
 PROSITE; PS00134; TRYPSIN_HIS; 1.
 PROSITE; PS00135; TRYPSIN_SER; 1.
 Plasminogen activation; Hydrolase; Serine protease; Glycoprotein; Signal.
 Plasma; Kringle; EGF-like domain; Repeat; Signal.
 PROPEP; 22 33 BY SIMILARITY.
 CHAIN; 34 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
 CHAIN; 34 314 TISSUE-TYPE PLASMINOGEN ACTIVATOR A CHAIN.
 CHAIN; 315 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR B CHAIN.
 DOMAIN; 40 82 FIBRONECTIN TYPE-1.
 DOMAIN; 83 121 EGF-LIKE.
 DOMAIN; 128 209 KRINGLE 1.
 DOMAIN; 219 300 KRINGLE 2.
 DOMAIN; 315 566 SERINE PROTEASE.
 ACT_SITE; 361 381 CHARGE RELAY SYSTEM.
 ACT_SITE; 410 410 CHARGE RELAY SYSTEM.
 ACT_SITE; 517 517 CHARGE RELAY SYSTEM.
 DISULFID; 42 72 BY SIMILARITY.
 DISULFID; 70 98 BY SIMILARITY.
 DISULFID; 92 109 BY SIMILARITY.
 DISULFID; 111 120 BY SIMILARITY.
 DISULFID; 128 209 BY SIMILARITY.
 DISULFID; 149 191 BY SIMILARITY.

FT DISULFID 180 204 BY SIMILARITY.
 FT DISULFID 219 300 BY SIMILARITY.
 FT DISULFID 240 282 BY SIMILARITY.
 FT DISULFID 271 295 BY SIMILARITY.
 FT DISULFID 303 434 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 346 362 BY SIMILARITY.
 FT DISULFID 354 423 BY SIMILARITY.
 FT DISULFID 448 523 BY SIMILARITY.
 FT DISULFID 480 496 BY SIMILARITY.
 FT DISULFID 513 541 BY SIMILARITY.
 FT CARBOHYD 153 153 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 487 487 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 566 AA; 63701 MW; 2E86EB4E32276C3 CRC64;
 Query Match 35.9%; Score 300.5; DB 1; Length 566;
 Best Local Similarity 44.0%; Pred. No. 8e-22;
 Matches 59; Conservative 16; Mismatches 54; Indels 5; Gaps 3;
 QY 3 ELHQVP-SNCD---CLNGGTCVSNKVFNSIHWNCNPKKGGQHCEDKSKTCYEGNGHFY 58
 DB 78 QCHSVPRVSCSEPCWFCNGGTCROALYSSDF-VCCPEGFMKCEIDATATCYKQGVAY 136
 QY 59 RGKASTDTMGRPCLPMNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYQVG 118
 DB 137 RGTWSTAESGAECANWSSSLAMKPYSGRPNALRLGLGNHNYCRNPDDQSKPCYVFK 196
 QY 119 LKPLVQECMVHDC 132
 DB 197 GYISEFCSTPACA 210
 RESULT 15
 ID -FA12_CAVPO STANDARD; PRT; 603 AA.
 AC Q04962;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor) (HAF) (Fragment).
 GN F12.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
 RC TISSUE=Liver;
 RX MEDLINE=93003367; PubMed=1390917;
 RA Samba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S., Kambata T., Okabe H.;
 RT "Primary structure of guinea-pig Hageman factor: sequence around the cleavage site differs from the human molecule."; Biochim. Biophys. Acta 1159:113-121(1992).
 CC -1- FUNCTION: Factor XII is a serum glycoprotein that participates in the initiation of blood coagulation, fibrinolysis, and the generation of bradykinin and angiotensin.
 CC -1- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-ile bonds in factor VII to form factor VIIa and factor XI to form factor Xla.
 CC -1- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a complex bound to an anionic surface. Prekallikrein is cleaved by factor XII to form kallikrein, which then cleaves factor XII first to alpha-factor Xlaa and then to beta-factor Xlaa. Alpha-factor Xlaa activates factor XI to factor Xla.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X68615; CAA48600.1; --
 HSRP; S28941; S28941.
 HSP; P00763; IDPO.
 MEROPS; S01.211; -- Cys_Ser_trypsin.
 InterPro; IPR009003; EGF_2.
 InterPro; IPR000742; EGF_2.
 InterPro; IPR001881; EGF_Ca.
 InterPro; IPR006209; EGF-like.
 InterPro; IPR000083; Fibnctnl.
 InterPro; IPR000582; FN_Type_II.
 InterPro; IPR006210; EGF.
 InterPro; IPR000001; Kringle.
 InterPro; IPR001254; Peptidase_S1.
 InterPro; IPR001314; Peptidase_S1A.
 Pfam; PF00008; EGF_2.
 Pfam; PF00039; fn1; 1.
 Pfam; PF00040; fn2; 1.
 Pfam; PF00051; kringle; 1.
 Pfam; PF00089; trypsin; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 PRINTS; PR00013; FNTYPEII.
 PRINTS; PR00018; KRINGLE.
 ProDom; PD000995; FN_Type_II; 1.
 ProDom; PD000395; Kringle; 1.
 SMART; SM00381; EGF; 2.
 SMART; SM00058; FN1; 1.
 SMART; SM00059; FN2; 1.
 SMART; SM00130; KR; 1.
 SMART; SM00020; Tryp_Spc; 1.
 PROSITE; PS00022; EGF_1; 2.
 PROSITE; PS01186; EGF_2; 1.
 PROSITE; PS00026; EGF_3; 2.
 PROSITE; PS01253; FIBRONECTIN_1; 1.
 PROSITE; PS00023; FIBRONECTIN_2; 1.
 PROSITE; PS00021; KRINGLE_1; 1.
 PROSITE; PS00070; KRINGLE_2; 1.
 PROSITE; PS00240; TRYPSIN_DOM; 1.
 PROSITE; PS00134; TRYPSIN_HIS; 1.
 PROSITE; PS00135; TRYPSIN_SER; 1.
 Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
 Hydrolyase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
 NON_TER 1
 SIGNAL <1 18
 CHAIN 19 358
 FT CHAIN 359 603
 DOMAIN 46 87
 DOMAIN 93 130
 DOMAIN 132 172
 DOMAIN 173 209
 DOMAIN 216 294
 DOMAIN 312 342
 DOMAIN 359 603
 DOMAIN 398 398
 ACT_SITE 447 447
 ACT_SITE 551 551
 DISULFID 97 109
 DISULFID 103 118
 DISULFID 120 129
 DISULFID 134 162
 DISULFID 160 169
 DISULFID 177 188
 DISULFID 182 197
 DISULFID 199 208
 DISULFID 216 294
 DISULFID 237 276
 DISULFID 265 289
 DISULFID 345 472
 DISULFID 383 399

FT DISULFID 391 461 BY SIMILARITY.
 FT DISULFID 422 425 BY SIMILARITY.
 FT DISULFID 488 557 BY SIMILARITY.
 FT DISULFID 520 536 BY SIMILARITY.
 FT DISULFID 547 578 BY SIMILARITY.
 FT CARBOHYD 248 248 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 270 270 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 419 419 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 603 AA; 66795 MW; 48DC6B946FB9ED59 CRC64;
 Query Match 32.0%; Score 268; DB 1; Length 603;
 Best Local Similarity 38.6%; Pred. No. 1.2e-18;
 Matches 51; Conservative 21; Mismatches 52; Indels 8; Gaps 3;
 QY 13 CLNGCTCVSNKYFSNIHWCNCPKFGQCHQCHIDKSKTCYEGNCHFYRGKASTDTMGPRCL 72
 Db 182 CLNGCRCL---VEGHLCDPCMGYTGPFCDLDTTASCYEGRGVSYRGARITVSGAKQ 238
 QY 73 PMSATVLQQTFAHRSD-ALQLGLGKHNYCRPNRRRRRRCVQVGLKPLVCECMVHDC 131
 Db 239 RWAS----EATYRMTAEALRRGLGHTTCRPNDRTRFCVFMGNRLSWEYCDLAQC 294
 QY 132 ADGKKSPSPPEE 143
 Db 295 QYPPQPTATPHD 306
 RESULT 16
 HGFA_HUMAN
 ID HGFA_HUMAN STANDARD; PRT; 655 AA.
 AC Q04756; Q14726;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
 DE activator) (HGFA).
 GN HGFA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Liver, and Serum;
 RX MEDLINE=93252878; PubMed=7683665;
 RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
 RA Kitamura N.;
 RT "Molecular cloning and sequence analysis of the cDNA for a human
 RT serine protease responsible for activation of hepatocyte growth
 RT factor. Structural similarity of the protease precursor to blood
 RT coagulation factor XII.";
 RL J. Biol. Chem. 268:10024-10028(1993).
 RN [2]
 RP SEQUENCE OF 40-655 FROM N.A.
 RA Zhao S., Odell C.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Activates hepatocyte growth factor (HGF) by
 CC converting it from a single chain to a heterodimeric form.
 CC -!- SUBUNIT: Dimer of a short chain and a long chain linked by a
 CC disulfide bond.
 CC -!- SUBCELLULAR LOCATION: Secreted as an inactive single-chain
 CC precursor and is then activated to a heterodimeric form.
 CC -!- TISSUE SPECIFICITY: Liver.
 CC -!- SIMILARITY: Belongs to peptidase family S1.
 CC -!- SIMILARITY: Contains 2 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- CAUTION: It is uncertain whether Met-1 is the initiator.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

EMBL; D14012; BAA03113.1; -; NOT_ANNOTATED_CDS.
EMBL; Z69923; -; NOT_ANNOTATED_CDS.
PIR; A46688; A46688.
HSSP; P00763; IDFO.
MEROPS; S01.228; -.
Genew; HGNC:4894; HGFAC.
MIM; 604552; -.
GO; GO:0005576; C:extracellular; TAS.
GO; GO:0004252; F:serine-type endopeptidase activity; TAS.
GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
InterPro; IPR009003; Cys_Ser_trypsin.
InterPro; IPR000742; EGF 2.
InterPro; IPR006209; EGF-like.
InterPro; IPR000083; Fibrinctnl.
InterPro; IPR000562; FN_Type_II.
InterPro; IPR006210; IEGF.
InterPro; IPR000001; Kringle.
InterPro; IPR001254; Peptidase_S1.
InterPro; IPR001314; Peptidase_S1A.
Pfam; PF00008; EGF; 2.
Pfam; PF00039; fnl; 1.
Pfam; PF00040; fn2; 1.
Pfam; PF00051; kringle; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00013; FNTYPEII.
PRINTS; PR00018; KRINGLE.
ProDom; PD000995; FN_Type_II; 1.
ProDom; PD000395; Kringle; 1.
SMART; SM00181; EGF; 2.
SMART; SM00058; FN1; 1.
SMART; SM00059; FN2; 1.
SMART; SM00130; KR; 1.
SMART; SM00020; TRYPSIN; 1.
PROSITE; PS00022; EGF 1; 2.
PROSITE; PS01186; EGF 2; 1.
PROSITE; PS00266; EGF 3; 2.
PROSITE; PS01253; FIBRONECTIN_1; 1.
PROSITE; PS00023; FIBRONECTIN_2; 1.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS00070; KRINGLE_2; 1.
PROSITE; PS00240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
EGF-like domain; Repeat; Zymogen.
SIGNAL 1 30
PROPEP 31 372
CHAIN 373 407
CLEAVED IN ACTIVE FORM.
HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT CHAIN
HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG CHAIN
CHAIN 408 655
DOMAIN 108 148
DOMAIN 160 198
DOMAIN 200 240
DOMAIN 241 279
DOMAIN 286 367
DOMAIN 408 655
ACT_SITE 447 447
ACT_SITE 497 497
ACT_SITE 598 598
DISULFID 108 133
DISULFID 122 148
DISULFID 164 175
DISULFID 169 186
DISULFID 188 197
DISULFID 202 230
DISULFID 228 237
FIBRONECTIN TYPE-II.
EGF-LIKE 1.
FIBRONECTIN TYPE-I.
EGF-LIKE 2.
KRINGLE
SERINE PROTEASE.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.

FT DISULFID 245. 256 BY SIMILARITY.
FT DISULFID 250 267 BY SIMILARITY.
FT DISULFID 269 278 BY SIMILARITY.
FT DISULFID 286 367 BY SIMILARITY.
FT DISULFID 307 349 BY SIMILARITY.
FT DISULFID 338 362 BY SIMILARITY.
FT DISULFID 394 521 INTERCHAIN (BY SIMILARITY).
FT DISULFID 432 448 BY SIMILARITY.
FT DISULFID 440 510 BY SIMILARITY.
FT DISULFID 535 604 BY SIMILARITY.
FT DISULFID 567 583 BY SIMILARITY.
FT DISULFID 594 622 BY SIMILARITY.
FT CARBOHYD 48 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 468 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 492 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 546 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 644 R -> Q (IN REF. 2).
SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1E1B862ED7 CRC64;
Query Match 31.3%; Score 262; DB 1; Length 655;
Best Local Similarity 36.9%; Pred. No. 5.1e-18;
Matches 58; Conservative 12; Mismatches 57; Indels 30; Gaps 3;
QY 5 HQVPSNCDCLNGGTCVSNKYFSNIHW-----CNCPRKFGQHCIDKSKTCYEGNGH 56
DB 242 HTACLSPCLNGGTC-----HLIVATGTVCACPGAGLGNIEPDERCFLENGT 292
QY 57 FYRGKASTDMGRPCLPWNSATVLTQTYHAHRSDALQLGKKNYCRNPNRRRPPWCYQ 116
DB 293 GYRGVASTASGLSCLAWNSDLLYQELHVDVSGAALLGLGHAYCRNPNDRPWCYV 352
QY 117 VGLKPLVQECMVHDC-----ADGKKFSPSP 140
DB 353 KDSALSWEYCRLEACESLRTVQLSPOLLATLPSPASP 389
RESULT 17
HGFA MOUSE
ID HGFA_MOUSE STANDARD; PRT; 653 AA.
AC QSR08; Q9JXV4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF activator) (HGFA).
GN HGFA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Itoh H., Kataoka H., Kono H.;
RT "Mouse hepatocyte growth factor activator.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=21226753; PubMed=11032833;
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J., Yang J., Huan Y.;
RT "Activation of hepatocyte growth factor (HGF) by endogenous HGF activator is required for mesangial kidney morphogenesis in vitro.";
RL J. Biol. Chem. 276:15099-15106(2001).
CC -!- FUNCTION: Activates hepatocyte growth factor (HGF) by converting it from a single chain to a heterodimeric form (By similarity).
CC -!- SUBUNIT: Dimer of a short chain and a long chain linked by a disulfide bond (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted as an inactive single-chain precursor and is then activated to a heterodimeric form (By similarity).

CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL; AF099017; AF02489.1; -;
 CC ENBL; AF224724; AAF34712.1; -;
 CC HSP; P00763; IDPO.
 CC MEROPS; S01.228; -;
 CC MGD; MG1.1859281; Hgfac.
 CC InterPro; IPR009003; Cys Ser trypsin.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR006209; EGF_like.
 CC InterPro; IPR000083; Fibnctn1.
 CC InterPro; IPR000562; FN_Type_II.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Peptidase S1.
 CC InterPro; IPR001314; Peptidase_S1A.
 CC Pfam; PF00008; EGF_2.
 CC Pfam; PF00039; fn1; 1.
 CC Pfam; PF00040; fn2; 1.
 CC Pfam; PF00051; kringle; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00013; PNTYPEII.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000995; FN_Type_II; 1.
 CC ProDom; PD000395; Kringle; 1.
 CC SMART; SM00181; EGF; 2.
 CC SMART; SM00059; FN2; 1.
 CC SMART; SM00130; KR; 1.
 CC SMART; SM00020; TYP_SPC; 1.
 CC PROSITE; PS00022; EGF_1; 2.
 CC PROSITE; PS01186; EGF_2; 1.
 CC PROSITE; PS00026; EGF_3; 2.
 CC PROSITE; PS01253; FIBRONECTIN_1; 1.
 CC PROSITE; PS00023; FIBRONECTIN_2; 1.
 CC PROSITE; PS00021; KRINGLE_1; 1.
 CC PROSITE; PS00021; KRINGLE_2; 1.
 CC PROSITE; PS50070; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
 KW EGF-like domain; Repeat; Zymogen.
 FT SIGNAL 1 29 BY SIMILARITY.
 FT PROPEP 30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).
 FT CHAIN 370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
 FT CHAIN
 FT CHAIN 406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG
 FT CHAIN
 FT DOMAIN 105 145 FIBRONECTIN TYPE-II.
 FT DOMAIN 157 195 EGF-LIKE 1.
 FT DOMAIN 197 237 FIBRONECTIN TYPE-I.
 FT DOMAIN 238 276 EGF-LIKE 2.
 FT DOMAIN 283 364 KRINGLE.
 FT DOMAIN 406 653 SERINE PROTEASE.
 FT ACT_SITE 445 445 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 596 596 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 105 130 BY SIMILARITY.
 FT DISULFID 119 145 BY SIMILARITY.
 FT DISULFID 161 172 BY SIMILARITY.
 FT DISULFID 166 183 BY SIMILARITY.

FT DISULFID 185 194 BY SIMILARITY.
 FT DISULFID 199 227 BY SIMILARITY.
 FT DISULFID 225 234 BY SIMILARITY.
 FT DISULFID 242 253 BY SIMILARITY.
 FT DISULFID 247 264 BY SIMILARITY.
 FT DISULFID 266 275 BY SIMILARITY.
 FT DISULFID 283 364 BY SIMILARITY.
 FT DISULFID 304 346 BY SIMILARITY.
 FT DISULFID 335 359 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 392 519 BY SIMILARITY.
 FT DISULFID 430 446 BY SIMILARITY.
 FT DISULFID 533 508 BY SIMILARITY.
 FT DISULFID 533 602 BY SIMILARITY.
 FT DISULFID 565 581 BY SIMILARITY.
 FT DISULFID 592 620 BY SIMILARITY.
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 164 164 G -> W (IN REF. 2).
 SQ SEQUENCE 653 AA; 70567 MW; 88B4B2025DF7FDC CRC64;
 Query Match 29.9%; Score 250.5; DB 1; Length 653;
 Best Local Similarity 42.0%; Pred. No. 6.6e-17;
 Matches 50; Conservative 11; Mismatches 41; Indels 17; Gaps 2;
 Qy 5 HQVPSNCDCLNGCTCYSNKYFSNIHW-----CNCPKFGGQHCEIDSKTCYSGNH 56
 Db 239 HTACLSFCLNGGTC-----HLIVGTGTSVCTCLGYAGRCNIVTEHCFLNGT 289
 Qy 57 FYRGKASTDTMGRCPLFPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPDRRRCYV 115
 Db 290 EYRGVASTAASGLSCLAWNSDLLVQELHVDVAVALGLGLPHAYCRNPKDRPCYV 348
 RESULT 18
 FA12 HUMAN
 ID FA12_HUMAN STANDARD; PRT; 615 AA.
 AC P00748; P78339;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
 DE (HAP).
 GN F12.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88007593; PubMed=2888762;
 RA Cool D.E., McGillivray R.T.A.;
 RT "Characterization of the human blood coagulation factor XII gene.
 RT Intron/exon gene organization and analysis of the 5'-flanking
 RT region."
 RL J. Biol. Chem. 262:13662-13673(1987).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605.
 RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
 RA Rajkumar N., Toch E.J., Yi Q., Nickerson D.A.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 4-615 FROM N.A.
 RX MEDLINE=86176794; PubMed=3754331;
 RA Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
 RA Cortese R.;
 RT "cDNA sequence coding for human coagulation factor XII (Hageman).";
 RL Nucleic Acids Res. 14:3146-3146(1986).
 RN [4]
 RP SEQUENCE OF 14-615 FROM N.A.

R MEDLINE=86033830; PubMed=3877053;
R Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
R McGilivray R.T.A.,
R "Characterization of human blood coagulation factor XII cDNA.
R Prediction of the primary structure of factor XII and the tertiary
R structure of beta-factor XIIa.";
R J. Biol. Chem. 260:13666-13676(1985).
R [5]
R SEQUENCE OF 146-615 FROM N.A.
R MEDLINE=86216049; PubMed=3011063;
R Que B.G., Davie E.W.,
R "Characterization of a cDNA coding for human factor XII (Hageman
R factor).";
R Biochemistry 25:1525-1528(1986).
R [6]
R SEQUENCE OF 20-379.
R MEDLINE=85182674; PubMed=3886654;
R McMullen B.A., Fujikawa K.,
R "Amino acid sequence of the heavy chain of human alpha-factor XIIa
R (activated Hageman factor).";
R J. Biol. Chem. 260:5328-5341(1985).
R [7]
R SEQUENCE OF 354-362 AND 373-615.
R MEDLINE=83291041; PubMed=6604055;
R Fujikawa K., McMullen B.A.,
R "Amino acid sequence of human beta-factor XIIa.";
R J. Biol. Chem. 258:10924-10933(1983).
R [8]
R SEQUENCE OF 561-615 FROM N.A.
R TISSUE=Blood;
R MEDLINE=96133302; PubMed=8528215;
R Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.,
R "The novel acceptor splice site mutation 11396(G->A) in the factor
R XII gene causes a truncated transcript in cross-reacting material
R negative patients.";
R Hum. Mol. Genet. 4:1235-1237(1995).
R [9]
R CARBOHYDRATE-LINKAGE SITE THR-109.
R MEDLINE=92184750; PubMed=1544894;
R Harris R.J., Ling V., Spellman M.W.,
R "O-linked fucose is present in the first epidermal growth factor
R domain of factor XII but not protein C.";
R J. Biol. Chem. 267:5102-5107(1992).
R [10]
R VARIANT WASHINGTON D.C. SER-590.
R MEDLINE=90046788; PubMed=2510163;
R Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B.,
R Saito H.,
R "Coagulation factor XII (Hageman factor) Washington D.C.: inactive
R factor XIIa results from Cys-571-->Ser substitution.";
R Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).
R [11]
R VARIANT LOCARNO PRO-372.
R MEDLINE=94325559; PubMed=8049433;
R Hovenga J.K., Schaller J., Stricker H., Willemin W.A., Furlan M.,
R Laemmle B.,
R "Coagulation factor XII Locarno: the functional defect is caused by
R the amino acid substitution Arg-353-->Pro leading to loss of a
R kallikrein cleavage site.";
R Blood 84:1173-1181(1994).
R [12]
R VARIANT TENRI CYS-53.
R MEDLINE=99290785; PubMed=10361128;
R Kondo S., Tokunaga F., Kawano S., Oono Y., Kumagai S., Koide T.,
R "Factor XII Tenri, a novel cross-reacting material negative factor XII
R deficiency, occurs through a proteasome-mediated degradation.";
R Blood 93:4300-4303(1999).
R C -!- FUNCTION: Factor XII is a serum glycoprotein that participates in
R the initiation of blood coagulation, fibrinolysis, and the
R generation of bradykinin and angiotensin.
R C -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
R VII to form factor VIIa and factor XI to form factor XIa.
R C -!- PTM: O- AND N-GLYCOSYLATED.

CC -!- DISEASE: Defects in F12 do not cause any clinical symptoms. The
CC sole effect is that whole-blood clotting time is prolonged.
CC -!- MISCELLANEOUS: Factor XII, prekallikrein, and HWW kininogen form a
CC complex bound to an anionic surface. Prekallikrein is cleaved by
CC factor XII to form kallikrein, which then cleaves factor XII first
CC to alpha-factor XIIa and then to beta-factor XIIa. Alpha-factor
CC XIIa activates factor XI to factor Xla.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M31315; AAA70225.1; -;
CC EMBL; AP538691; AAM97932.1; -;
CC EMBL; M11723; AA51986.1; -;
CC EMBL; M17466; AAB59490.1; -;
CC EMBL; M17464; AAB59490.1; JOINED.
CC EMBL; M17465; AAB59490.1; JOINED.
CC EMBL; M13147; AAA70224.1; -;
CC EMBL; U71274; AAB51203.1; -;
CC PIR; A29411; KFHU12.
CC HSP; P00763; IDPO.
CC MEROXS; S01.211; -;
CC Genes; HGNC:3530; F12.
CC MIM; 234000; -;
CC GO; GO:0003805; F12 blood coagulation factor XI activity; TAS.
CC GO; GO:0003806; F12 blood coagulation factor XII activity; TAS.
CC GO; GO:0008236; F12 serine-type peptidase activity; TAS.
CC GO; GO:0007596; F12 blood coagulation; TAS.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR062059; EGF-like.
CC InterPro; IPR000083; Fibrinctnl.
CC InterPro; IPR00562; FN_Type_II.
CC InterPro; IPR06210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00008; EGF_2.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00040; fn2; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00013; FNTYPEII.
CC PRINTS; PR00016; KRINGLE.
CC ProDom; PD000995; FN_Type_II; 1.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF_2.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00059; FN2; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS00026; EGF_3; 2.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00023; FIBRONECTIN_2; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hydratase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;

F12.
Bos taurus (Bovine).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
[1]
SEQUENCE FROM N.A.
TISSUE=Liver;
MEDLINE=94242782; PubMed=8186251;
Shibuya Y., Semba U., Okabe H., Kambara T., Yamamoto T.;
"Primary structure of bovine Hageman factor (blood coagulation factor
XII): comparison with human and guinea pig molecules.";
Biochim. Biophys. Acta 1206:63-70 (1994).
[2]
SEQUENCE OF 10-21; 350-364 And 525-550.
MEDLINE=77182112; PubMed=861210;
Fujikawa K., Walsh A.K., Davie W.E.;
"Isolation and characterization of bovine factor XII (Hageman
factor).";
Biochemistry 16:2270-2278 (1977).
-!- FUNCTION: Factor XII is a serum glycoprotein that participates in
the initiation of blood coagulation, fibrinolysis, and the
generation of bradykinin and angiotensin.
-!- CATALYTIC ACTIVITY: Cleaves selectively Arg-Ile bonds in factor
XII to form factor VIIa and factor XI to form factor Xla.
-!- PTM: O- AND N-Glycosylated (BY SIMILARITY).
-!- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a
complex bound to an anionic surface. Prekallikrein is cleaved by
factor XII to form kallikrein, which then cleaves factor XII first
to alpha-factor Xlla and then to beta-factor Xlla. Alpha-factor
Xlla activates factor XI to factor Xla. Bovine factor XII is
cleaved only to alpha-factor Xlla as it lacks the trypsin/
kallikrein cleavage site.
-!- SIMILARITY: Belongs to peptidase family S1.
-!- SIMILARITY: Contains 2 EGF-like domains.
-!- SIMILARITY: Contains 1 fibronectin type I domain.
-!- SIMILARITY: Contains 1 fibronectin type II domain.
-!- SIMILARITY: Contains 1 kringle domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; S70164; AAB30804.2; -;
PIR; S45281; S45281.
HSP; P00763; IDPO.
MEROPS; S01.211; -;
InterPro; IPR009003; Cys Ser trypsin.
InterPro; IPR006209; EGF-like.
InterPro; IPR000083; Fibnctnl.
InterPro; IPR000562; FN Type_II.
InterPro; IPR006210; IEGF.
InterPro; IPR000001; Kringle.
InterPro; IPR001254; Peptidase S1.
InterPro; IPR001314; Peptidase_S1A.
Pfam; PF00008; EGF; 2.
Pfam; PF00039; fn1; 1.
Pfam; PF00040; fn2; 1.
Pfam; PF00051; kringle; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00013; FNTYPEII.
PRINTS; PR00018; KRINGLE.
ProDom; P000995; FN_Type_II; 1.
ProDom; P000395; Kringle; 1.
SMART; SM00181; EGF; 2.
SMART; SM00058; FN1; 1.
SMART; SM00059; FN2; 1.

DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hydratase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
FT NON_TER 1 1
FT SIGNAL <1 9
FT CHAIN 10 349
FT CHAIN 350 593
FT DOMAIN 37 78
FT DOMAIN 84 121
FT DOMAIN 123 163
FT DOMAIN 164 200
FT DOMAIN 207 287
FT DOMAIN 297 333
FT DOMAIN 350 593
FT ACT_SITE 389 389
FT ACT_SITE 438 438
FT ACT_SITE 541 541
FT DISULFID 88 100
FT DISULFID 94 109
FT DISULFID 111 120
FT DISULFID 125 133
FT DISULFID 151 160
FT DISULFID 168 179
FT DISULFID 173 188
FT DISULFID 190 199
FT DISULFID 207 287
FT DISULFID 230 289
FT DISULFID 258 282
FT DISULFID 336 463
FT DISULFID 374 390
FT DISULFID 382 452
FT DISULFID 413 416
FT DISULFID 479 547
FT DISULFID 510 526
FT DISULFID 537 588
FT CARBOHYD 99 99
FT CARBOHYD 241 241
FT CARBOHYD 263 263
FT CARBOHYD 410 410
SQ SEQUENCE 593 AA; 65148 MW; 721592BA792BD61F CRC64;

Query Match 25.9%; Score 217; DB 1; Length 593;
Best Local Similarity 35.7%; Pred. No. 1.1e-13;
Matches 46; Conservative 15; Mismatches 58; Indels 10; Gaps 4;

QY 6 QVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCVE--GNGHFYRGKAS 63
DB 166 QVCRTPNCLNGSDCLQAE---GRLCRCAPSPAGRLCDVDLXASCYDDRDGLSRGMAG 222
QY 64 TDTMGRPCLPWNASATVLOQTY-HAHRSDALQGLGKHNYCRNPDRRRPWCYQVGLKPL 122
DB 223 TTLSGAPCQSWAS-----EATYWNVTAEQVLNWLGLGHAFCRPNPDNDRPWCIFWKGDRLS 278
QY 123 VQECMVHDC 131
DB 279 WNYCRLAPC 287

RESULT 21
RORI DROME
ID RORI DROME
AC Q24488; STANDARD; PRT; 685 AA.

28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor Ror precursor
DE (EC 2.7.1.112) (dRor).
GN ROR OR CG4926.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Spheroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=Canton-S; TISSUE=Larval brain;
RX MEDLINE=93348222; PubMed=8394009;
RA Wilson C., Gohardian D.C.I., Steller H.;
RT "Ror", a potential neurotrophic receptor gene, encodes a Drosophila
RT homolog of the vertebrate Ror family of trk-related receptor tyrosine
RT kinases.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:7109-7113(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196066; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernier B.P., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flosser C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heilmann T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "the genome sequence of Drosophila melanogaster.";
RL Science 287:12185-12195(2000).
RN [3]
RP SEQUENCE OF 545-597 FROM N.A.
RX MEDLINE=98401146; PubMed=9731193;
RA Oates A.C., Wollberg P., Achen M.G., Wilks A.F.;
RT "Sampling the genomic pool of protein tyrosine kinase genes using the
RT polymerase chain reaction with genomic DNA".
RL Biochem. Biophys. Res. Commun. 249:660-667(1998).
CC -!- FUNCTION: Tyrosine-protein kinase receptor that functions during
CC early stages of neuronal development.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.

-!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
-!- TISSUE SPECIFICITY: Expressed in neurons of the developing nervous
CC system.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR
CC subfamily.
CC -!- SIMILARITY: Contains 1 frizzled (Fz) domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L20297; AAA28860.1; -;
CC EMBL: AB003628; AAF52855.1; -;
CC EMBL: AJ002908; CAA05743.1; -;
CC PIR: A48289; A48289.
CC HSSP: P11362; 1FGK.
CC FlyBase: FBgn0010407; Ror.
CC GO: GO:0016021; C: integral to membrane; NAS.
CC GO: GO:0004713; F: protein-tyrosine kinase activity; NAS.
CC GO: GO:0007417; P: central nervous system development; IEP.
CC GO: GO:0006468; P: protein amino acid phosphorylation; NAS.
CC InterPro: IPR000024; Fz domain.
CC InterPro: IPR000001; Kringle.
CC InterPro: IPR000719; Prot kinase.
CC InterPro: IPR002011; RecepttyrkinII.
CC InterPro: IPR001245; Tyr_kinase_AS.
CC InterPro: IPR008266; Tyr_kinase_AS.
CC Pfam: PF00051; Kringle; 1.
CC Pfam: PF00069; Kkinase; 1.
CC PRINTS: PR00109; TYRKINASE.
CC PRINTS: PR00109; TYRKINASE.
CC ProDom: PD000395; Kringle; 1.
CC ProDom: PD000001; Prot_kinase; 1.
CC SMART: SM00130; KR_1.1.
CC SMART: SM00219; TYRK; 1.
CC PROSITE: PS00038; FZ; 1.
CC PROSITE: PS00021; KRINGLE_1; 1.
CC PROSITE: PS00070; KRINGLE_2; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00103; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS00239; RECEPTOR_TYR_KIN_II; 1.
CC Transferase; Kinase; Tyrosine-protein kinase; ATP-binding; Receptor;
KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
KW Developmental protein.
FT SIGNAL 1 24
FT CHAIN 25 685
FT TYROSINE-PROTEIN KINASE TRANSMEMBRANE
FT RECEPTOR ROR (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT FZ.
FT KRINGLE.
FT PROTEIN KINASE.
FT ATP (BY SIMILARITY).
FT ATP (BY SIMILARITY).
FT BY SIMILARITY.
FT PHOSPHORYLATION (AUTO-) (BY
FT SIMILARITY).
FT PHOSPHORYLATION (AUTO-) (BY
FT SIMILARITY).
FT PHOSPHORYLATION (AUTO-) (BY
FT SIMILARITY).
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 45
FT CARBOHYD 63
FT CARBOHYD 129
FT CARBOHYD 144
FT CARBOHYD 144
FT CARBOHYD 250

S SEQUENCE 685 AA; 78142 MW; 526162D27D5DPD7C7 CRC64;
Very Match 20.2%; Score 169; DB 1; Length 685;
Local Similarity 28.8%; Pred. No. 5.7e-09;
atches 38; Conservative 27; Mismatches 35; Indels 32; Gaps 6;
Q 3 ELHVPSCDCLNGTCTVSNKFSNIHWCNPKKFGGHCIDKSKTCYEGNGHFYRGA 62
D 212 DCKLPQKHCLSLGITI-----EVDKENCYWEDGSTYRGVA 249
Q 63 SDTMRGRLPNSAVLQOTYVHRSALQGLGKHNYCRNPON-RRRPWCYVOVGLKP 121
D 250 NVSAGGKPLRN-SWLKKEI-----SDPEL-IGQ-NYCRNPGSVENSPWCFVDSRRER 300
Q 122 LVQECWHDCA 133
D 301 IIECDIPKCA 312
R JLT 22
A HUMAN
I APOA HUMAN STANDARD; PRT; 4548 AA.
A P08519;
D 01-AUG-1988 (Rel. 08, Created)
D 01-AUG-1988 (Rel. 08, Last sequence update)
D 10-OCT-2003 (Rel. 42, Last annotation update)
D Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).
C LPA.
O Homo sapiens (Human).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
C NCBI_TaxID:9606;
K [1]
R SEQUENCE FROM N.A.
E MEDLINE=8903109; PubMed=3670400;
R McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,
R Fless G.M., Scannu A.M., Lawn R.M.;
R "cdna sequence of human apolipoprotein(a) is homologous to
R plasminogen.";
R Nature 330:132-137(1987).
R [2]
R SERINE PROTEASE ACTIVITY.
R MEDLINE=90076123; PubMed=2531657;
R Salonen E.-M., Jauhainen M., Zardi L., Vaheri A., Ehnholm C.;
R "Lipoprotein(a) binds to fibronectin and has serine proteinase
R activity capable of cleaving it.";
R EMBO J. 8:4035-4040(1989).
R [3]
R REVIEW.
R MEDLINE=90049223; PubMed=2530631;
R Utermann G.;
R "The mysteries of lipoprotein(a).";
R Science 246:904-910(1989).
R [4]
R CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.
R MEDLINE=21303595; PubMed=11294842;
R Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;
R "Structural elucidation of the N- and O-glycans of human
R apolipoprotein(a): role of o-glycans in conferring protease
R resistance.";
R J. Biol. Chem. 276:22200-22208(2001).
R [5]
R X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.
R MEDLINE=96217891; PubMed=8642595;
R Nikol V., Locrasso P.V., Boettcher B.R.;
R "Crystal structures of apolipoprotein(a) kringle IV37 free and
R complexed with 6-aminohexanoic acid and with p-aminomethylbenzoic
R acid: existence of novel and expected binding modes.";
R J. Mol. Biol. 256:751-761(1996).
R [6]
R VARIANT ARG-4193.
R MEDLINE=95002201; PubMed=7918682;
R Scannu A.N., Pfaffinger D., Lee J.C., Hinman J.;
RT "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37
RT associated with a lysine binding defect in Lp(a).";
RL Biochim. Biophys. Acta 1227:41-45(1994).
CC -I- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.
CC -I- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin.
CC -I- PTM: N- and O-glycosylated. The N-glycans are complex biantennary
CC structures present in either a mono- or disialylated state. The
CC O-glycans are mostly (80%) represented by the monosialylated core
CC type I structure, NeuAcalpha2-3Galbeta1-3GalNAc, with smaller
CC amounts of disialylated and non-sialylated O-glycans also
CC detected.
CC -I- DISEASE: Elevated plasma concentrations of apo(a) and its
CC naturally occurring proteolytic fragments are correlated with
CC atherosclerosis. Homology with plasminogen kringles IV and V is
CC thought to underlie the atherogenicity of the protein, because the
CC fragments are competing with plasminogen for fibrin(ogen) binding.
CC -I- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC leading to the formation of the so called mini-Lp(a). Apo(a)
CC fragments accumulate in atherosclerotic lesions, where they may
CC promote thrombogenesis. O-glycosylation may limit the extent of
CC proteolytic fragmentation.
CC -I- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -I- SIMILARITY: Contains 38 kringle domains.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X06290; CAA29618.1; -
CC PIR: S00657; S00657.
CC PDB: 1I71; 13-JUN-01.
CC PDB: 1JFN; 28-JUN-02.
CC PDB: 1KIV; 18-MAY-99.
CC PDB: 3KIV; 18-MAY-99.
CC PDB: 4KIV; 18-MAY-99.
CC MEROPS: S01.226; -
CC Genew, HGNC:6667; LPA.
CC Min; 152200; -
CC GO: GO:0004866; F: endopeptidase inhibitor activity; TAS.
CC GO: GO:0008015; P: circulation; TAS.
CC GO: GO:0009405; P: pathogenesis; TAS.
CC InterPro: IPR009003; Cys Ser trypsin.
CC InterPro: IPR000001; Kringle.
CC InterPro: IPR001254; Peptidase S1.
CC InterPro: IPR001314; Peptidase_S1A.
CC Pfam: PF00051; Kringle; 38.
CC PRINTS: PR00089; trypsin; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC PRINTS: PR00018; KRINGLE.
CC ProDom: PD000395; Kringle; 38.
CC SMART: SM00130; KR; 38.
CC SMART: SM00020; TrypSPc; 1.
CC PROSITE: PS00021; KRINGLE_1; 38.
CC PROSITE: PS00070; KRINGLE_2; 38.
CC PROSITE: PS0240; TRYPsin DOM; 1.
CC PROSITE: PS00134; TRYPsin HIS; 1.
CC PROSITE: PS00135; TRYPsin SER; 1.
CC Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW Kringle; Repeat; Atherosclerosis; Signal; Polymorphism; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 4548 APOLIPOPROTEIN(A).
FT DOMAIN 20 130 KRINGLE TYPE IV, 1.
FT DOMAIN 131 244 KRINGLE TYPE IV, 2.
FT DOMAIN 245 358 KRINGLE TYPE IV, 3.
FT DOMAIN 359 472 KRINGLE TYPE IV, 4.

FT DOMAIN 473 586 KRINGLE TYPE IV, 5.
FT DOMAIN 587 700 KRINGLE TYPE IV, 6.
FT DOMAIN 701 814 KRINGLE TYPE IV, 7.
FT DOMAIN 815 928 KRINGLE TYPE IV, 8.
FT DOMAIN 929 1042 KRINGLE TYPE IV, 9.
FT DOMAIN 1043 1156 KRINGLE TYPE IV, 10.
FT DOMAIN 1157 1270 KRINGLE TYPE IV, 11.
FT DOMAIN 1271 1384 KRINGLE TYPE IV, 12.
FT DOMAIN 1385 1498 KRINGLE TYPE IV, 13.
FT DOMAIN 1499 1612 KRINGLE TYPE IV, 14.
FT DOMAIN 1613 1726 KRINGLE TYPE IV, 15.
FT DOMAIN 1727 1840 KRINGLE TYPE IV, 16.
FT DOMAIN 1841 1954 KRINGLE TYPE IV, 17.
FT DOMAIN 1955 2068 KRINGLE TYPE IV, 18.
FT DOMAIN 2069 2182 KRINGLE TYPE IV, 19.
FT DOMAIN 2183 2297 KRINGLE TYPE IV, 20.
FT DOMAIN 2297 2410 KRINGLE TYPE IV, 21.
FT DOMAIN 2411 2524 KRINGLE TYPE IV, 22.
FT DOMAIN 2525 2638 KRINGLE TYPE IV, 23.
FT DOMAIN 2639 2752 KRINGLE TYPE IV, 24.
FT DOMAIN 2753 2866 KRINGLE TYPE IV, 25.
FT DOMAIN 2867 2980 KRINGLE TYPE IV, 26.
FT DOMAIN 2981 3094 KRINGLE TYPE IV, 27.
FT DOMAIN 3095 3208 KRINGLE TYPE IV, 28.
FT DOMAIN 3209 3322 KRINGLE TYPE IV, 29.
FT DOMAIN 3323 3436 KRINGLE TYPE IV, 30.
FT DOMAIN 3437 3550 KRINGLE TYPE IV, 31.
FT DOMAIN 3551 3664 KRINGLE TYPE IV, 32.
FT DOMAIN 3665 3778 KRINGLE TYPE IV, 33.
FT DOMAIN 3779 3894 KRINGLE TYPE IV, 34.
FT DOMAIN 3895 3998 KRINGLE TYPE IV, 35.
FT DOMAIN 3999 4112 KRINGLE TYPE IV, 36.
FT DOMAIN 4113 4226 KRINGLE TYPE IV, 37.
FT DOMAIN 4227 4340 KRINGLE TYPE V.
FT ACT SITE 4341 4454 SERINE PROTEASE.
FT ACT SITE 4455 4568 CHARGE RELAY SYSTEM.
FT ACT SITE 4569 4682 CHARGE RELAY SYSTEM.
FT ACT SITE 4683 4796 CHARGE RELAY SYSTEM.
FT VARIAT 4797 4910 W -> R (LOSS OF LYSINE-SEPHAROSE BINDING).
FT SEQUENCE 4911 5024 /FTIG=VAR 006633.
FT SEQUENCE 5025 5138 501313 MW; 96921BE96A465C5F CRC64;
FT Query Match 19.4%; Score 162.5; DB 1; Length 4548;
FT Best Local Similarity 33.1%; Pred. No. 1.6e-07;
FT Matches 41; Conservative 12; Mismatches 46; Indels 25; Gaps 7;
QY 27 NIHW--CN---CP-----KFGGQRCEDKSKT---CYEGNGHFYRGKASTDTMGR 69
Db 3742 NVREWCNLTQCPVTSSVLATSTAVSECAPEQSPVTQDCYHGDGQSVRGFSFTVTGR 3801
QY 70 PCLPWSATVLOQTYHAHSDALQLGLKHNCRPNRRPWCYQVGLKPLV--QECM 127
Db 3802 TCOSWSSMT---PHWHQRTTEYPNGGLTRNCRNPDAEIREWCYT--MDPSVRWEYCN 3855
QY 128 VHDC131
Db 3856 LTQC13859
RE IT 23
KR RAT
ID KRM1 RAT STANDARD; PRT; 473 AA.
AC Q92434;
DI 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (kringle-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1 OR KREMEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OK NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakamura T., Nakamura T.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -i- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity)(Potential).
CC -i- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -i- SIMILARITY: Contains 1 CUB domain.
CC -i- SIMILARITY: Contains 1 kringle domain.
CC -i- SIMILARITY: Contains 1 WSC domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB065090; BAB62003.1; -
DR GO: GO:0016021; C-integral to membrane; ISS.
DR InterPro: IPR000859; CUB.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR002889; WSC.
DR Pfam: PF00431; CUB; 1.
DR Pfam: PF00051; kringle; 1.
DR PRINTS: PR001822; WSC; 1.
DR PRINTS: PR00018; KRINGLE.
DR PRODOM: PD000395; Kringle; 1.
DR SMART: SM00042; CUB; 1.
DR SMART: SM00130; KR; 1.
DR PROSITE: PS01180; CUB; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
KW Wnt signaling pathway; Signal; Transmembrane; Kringle.
FT SIGNAL 1 19
FT CHAIN 20 473 KREMEN PROTEIN 1.
FT DOMAIN 21 392 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 393 413 POTENTIAL.
FT DOMAIN 414 473 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 114 KRINGLE.
FT DOMAIN 120 210 WSC.
FT DOMAIN 214 321 CUB.
FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 473 AA; 51869 MW; 98510857DF856F08 CRC64;
Query Match 19.2%; Score 161; DB 1; Length 473;
Best Local Similarity 45.1%; Pred. No. 2.3e-08;
Matches 32; Conservative 7; Mismatches 28; Indels 4; Gaps 2;
QY 47 SKTCYSGNGHFYRGKASTDTM--GRCLPWSATVLOQTYHAHSDALQLGLKHNCRN 104
Db 29 SPECFTANGADYRGTSWTALQGGKCLFWNE--TFQHPYNTLKYNGEGGLGEHNYCRN 86
QY 105 PDNRREPWCYV 115
Db 87 PDGVSFPCYV 97
RESULT 24
APOA_MACMU
ID APOA_MACMU STANDARD; PRT; 1420 AA.
AC P14417;
DI 01-JAN-1990 (Rel. 13, Created)
DI 01-JAN-1990 (Rel. 13, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (lp(a)) (Fragment).

LPA.
Macaca mulatta (Rhesus macaque).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
NCBI_TaxID=9544;
[1]
SEQUENCE FROM N.A.
MEDLINE=89174660; PubMed=2925643;
Tomlinson J.E., McLean J.W., Lawn R.M.;
"Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
synthesis".
J. Biol. Chem. 264:5957-5965(1989).
-!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
(Lp(a)). It has serine proteinase activity and is able of
autocatalysis. Inhibits tissue-type plasminogen activator 1.
Lp(a) may be a ligand for megalin/Gp 330.
-!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
decorin (By similarity).
-!- PTM: N- and O-glycosylated (By similarity).
-!- DISEASE: Elevated plasma concentrations of apo(a) and its
naturally occurring proteolytic fragments are correlated with
atherosclerosis. Homology with plasminogen kringle IV and V is
thought to underlie the atherogenicity of the protein, because the
fragments are competing with plasminogen for fibrinogen binding.
-!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
leading to the formation of the so called mini-Lp(a). Apo(a),
fragments accumulate in atherosclerotic lesions, where they may
promote thrombogenesis. O-glycosylation may limit the extent of
proteolytic fragmentation (By similarity).
-!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
-!- SIMILARITY: Contains at least 10 kringle domains.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; J04635; AAA36933.1; -
PIR; A32869; A32869.
HSSP; P00747; 2PK4.
MEROPS; S01.226; -
InterPro; IPR009003; Cys Ser trypsin.
InterPro; IPR000001; Kringle_
InterPro; IPR001254; Peptidase S1.
InterPro; IPR001314; Peptidase_S1A.
Pfam; PF00051; kringle; 11.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; KRINGLE.
ProDom; PD000395; Kringle; 10.
SMART; SM00130; KR; 10.
SMART; SM00020; Tryp_Src; 1.
PROSITE; PS00021; KRINGLE_1; 10.
PROSITE; PS00070; KRINGLE_2; 10.
PROSITE; PS00240; TRYPsin_DOM; 1.
PROSITE; PS00134; TRYPsin_HIS; FALSE NEG.
PROSITE; PS00135; TRYPsin_SER; FALSE NEG.
Kl
Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
Kl
Klingle; Repeat; Atherosclerosis.
NON_TER 1
DOMAIN 1
DOMAIN 49 127 KRINGLE 1.
DOMAIN 163 241 KRINGLE 2.
DOMAIN 277 355 KRINGLE 3.
DOMAIN 391 469 KRINGLE 4.
DOMAIN 505 583 KRINGLE 5.
DOMAIN 619 697 KRINGLE 6.
DOMAIN 725 803 KRINGLE 7.
DOMAIN 839 917 KRINGLE 8.
DOMAIN 953 1031 KRINGLE 9.

FT DOMAIN 1067 1145 KRINGLE 10.
FT DOMAIN 1191 1420 SERINE PROTEASE.
SQ SEQUENCE 1420 AA; 158367 MW; BE102949E03C5BOE CRC64;
Query Match 19.2%; Score 161; DB 1; Length 1420;
Best Local Similarity 40.9%; Pred. No. 7e-08;
Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;
QY 50 CYEGNGHYFYRGKASDTMGPRCLPWNSATVLC--QTYHAHRSDALQGLGKKNYCNPN 107
Db 1068 CYHGNGQSYRGFTVITVGTGTCQSSMTFPHQKTFPENHPNDLTM-----NYCENPDA 1122
QY 108 RRPWCYVQVGLKPLVQE--CMVHDCAD 133
Db 1123 DTGPWCFT---MDPSVREYCNLTRCS 1147
RESULT 25
PLMN MACMU
ID PLMN MACMU STANDARD; PRT; 810 AA.
AC P12545;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7).
GN PLG.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89174660; PubMed=2925643;
RA Tomlinson J.E., McLean J.W., Lawn R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
synthesis".
RL J. Biol. Chem. 264:5957-5965(1989).
CC -!- FUNCTION: plasmin dissolves the fibrin of blood clots and acts as
a proteolytic factor in a variety of other processes including
embryonic development, tissue remodeling, tumor invasion, and
inflammation; in ovulation it weakens the walls of the Graafian
follicle. It activates the urokinase-type plasminogen activator,
collagenases and several complement zymogens, such as C1 and C5.
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von
Willebrand factor.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
products.
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen
CC activators, both plasminogen and its activator being bound to
CC fibrin. Activated with catalytic amounts of streptokinase.
CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
CC immediately after dissociation from the clot.
CC -!- MISCELLANEOUS: In the presence of the inhibitor, the activation
CC involves only cleavage after Arg-580, resulting in 2 chains held
CC together by 2 disulfide bonds. Without the inhibitor, the
CC activation involves also removal of the activation peptide.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -!- SIMILARITY: Contains 5 kringle domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; J04697; AAA36901.1; -
DR PIR; B32869; B30848.
DR HSSP; P00747; 1PMK.
DR MEROPS; S01.233; -.

```

Query Match          19.1%; Score 159.5; DB 1; Length 810;
Best Local Similarity 31.4%; Pred. No. 5.6e-08;
Matches 49; Conservative 12; Mismatches 66; Indels 29; Gaps 7;

QY      10  NCDCLNGGTCVSNKYFNSNIHWNCNPKKFGGQ-----HCEIDKSKTCYEGNGH 56
Db      428  NPDADKGPWCFTDPSVRWEYCNLKCSCTEGSAVAPPVPAQLPDAETPSEEDCMFGNGK 487

QY      57  FVRGKASTDTMGRCPLPWNASATVLQCTYHAHR-----SDALGLGKHNYCRNPD-NRRRP 111
Db      488  GRRGKATVTGTGFCQEWAA-----QEPHSIRIFTETNPRAGLEK-NYCRNPDGVDGVP 541

QY      112  WCYVQVGLKPLVQECMVHDCA----DGKPSPPPEE 143
Db      542  WCYT-TNPRKLFYCDVPQCAASSFDCGKQPVEPKK 576

RESULT 26
ID      KRL1_XENLA      STANDARD;      PRT;      452 AA.
AC      Q90Y30;
DT      10-OCT-2003 (Rel. 42, Created)
DT      10-OCT-2003 (Rel. 42, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Kremen protein 1 precursor (Kringle-containing protein marking the eye
DE      and the nose) (Dickkopf receptor).
GN      KREMEN1.
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC      Xenopodinae; Xenopus.
ON      NCBI_TaxID=8355;
RX      [1]
RP      SEQUENCE FROM N.A.
RA      Nakamura T.;
RL      Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: Receptor for dickkopf protein. Cooperates with Dickkopf
CC      to block Wnt/beta-catenin signaling (By similarity).
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC      -1- SIMILARITY: Contains 1 CUB domain.
CC      -1- SIMILARITY: Contains 1 kringle domain.
CC      -1- SIMILARITY: Contains 1 WSC domain.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AB070851; BAB64294.1; -.
CC      InterPro; IPR000859; CUB.
CC      InterPro; IPR000001; Kringle.
CC      InterPro; IPR002889; WSC.
CC      Pfam; PF00431; CUB; 1.
CC      Pfam; PF00051; kringle; 1.
CC      Pfam; PF01822; WSC; 1.
CC      PRINTS; PR00018; KRINGLE.
CC      ProDom; PD000395; Kringle; 1.
CC      SMART; SMC0042; CUB; 1.
CC      SMART; SMC0130; KR; 1.
CC      SMART; SMC0321; WSC; 1.
CC      PROSITE; PS01180; CUB; 1.
CC      PROSITE; PS00021; KRINGLE 1; 1.
CC      PROSITE; PS50070; KRINGLE-2; 1.
KW      Wnt signaling pathway; Glycoprotein; Kringle;
KW      Transmembrane; Signal.
FT      SIGNAL          1..22
FT      CHAIN           23..452
FT      DOMAIN          23..369
FT      TRANSMEM       370..390
FT      POTENTIAL
FT      EXTRACELLULAR (POTENTIAL).

```

F DOMAIN 331 452 CYTOPLASMIC (POTENTIAL).
F DOMAIN 29 112 KRINGLE.
F DOMAIN 118 208 WSC.
F DOMAIN 212 319 CUB.
F CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
F CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
F CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
F CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
F CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
F CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
F CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
S SEQUENCE 452 AA; 50188 MW; ED24BCD1AP4564E2 CRC64;

Query Match 19.0%; Score 159; DB 1; Length 452;
Best Local Similarity 40.8%; Pred. No. 3.5e-08;
Matches 31; Conservative 9; Mismatches 32; Indels 4; Gaps 2;

Q 42 CEIDSKTCYEGNGHYFGKASTDTM--GRCLPWSATVLQOQTYHAHRSDALQGLGKH 99
D 22 CSDSFHSECYTVNGADYRGTSQNTQTSIDGKPCFLWNE--TFQHPYNTLKYPNGGGLGEH 79

Q 100 NYCRNPDNRPPWCYV 115
D 80 NYCRNPDGVDSPWCYI 95

R JLT 27
K MOUSE STANDARD; PRT; 473 AA.
I KRMJ_MOUSE STANDARD; PRT; 473 AA.
A Q99N43; 28-FEB-2003 (Rel. 41, Created)
D 28-FEB-2003 (Rel. 41, Last sequence update)
D 10-OCT-2003 (Rel. 42, Last annotation update)
D Kremen protein 1 precursor (Kringle-containing protein marking the eye and the nose) (Dickkopf receptor).
D KREMN1 OR KREMEN.
C Mus musculus (Mouse).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
C NCBI_TaxID=10090;
K [1] SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.
E TISSUE=Brain, and Kidney;
R MEDLINE=21167372; PubMed=11267660;
K Nakamura T., Aoki S., Kitajima K., Takahashi T., Matsumoto K., Nakamura T.;
R "Molecular cloning and characterization of Kremen, a novel kringle-containing transmembrane protein.";
R Biochim. Biophys. Acta 1518:63-72(2001).
C -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf to block Wnt/beta-catenin signaling (By similarity).
C -!- SUBCELLULAR LOCATION: Type I membrane protein (potential).
C -!- TISSUE SPECIFICITY: In the adult, widely expressed with high levels in heart, lung, kidney, skeletal muscle and testis.
C -!- DEVELOPMENTAL STAGE: In the embryo, expression is first detected on day 9 and increases up to day 18. Lower levels are found in adult. At 9.5 dpc, expression is localised to the apical ectodermal ridge (AR) of the developing fore- and hindlimb buds, the telencephalon and the first brachial arch. At 10.5 dpc, expression is also observed in the myotome and in sensory tissues such as the nasal pit and optic vesicle.
C -!- SIMILARITY: Contains 1 CUB domain.
C -!- SIMILARITY: Contains 1 kringle domain.
C -!- SIMILARITY: Contains 1 WSC domain.

C This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AB059617; BAB40968.1; -.
DR HSP; P00747; ICEA.
DR MGD; MGI:1933988; Kremen.
DR GO; GO:0016021; C:Integral to membrane; NAS.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR002893; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
KW Wnt signaling pathway; Signal; Transmembrane; Kringle.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 473 KREMEN PROTEIN 1.
FT DOMAIN 21 392 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 393 413 POTENTIAL.
FT DOMAIN 414 473 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 114 KRINGLE.
FT DOMAIN 120 210 WSC.
FT DOMAIN 214 321 CUB.
FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 473 AA; 51716 MW; 586827788BE3FDD1 CRC64;

Query Match 19.0%; Score 159; DB 1; Length 473;
Best Local Similarity 45.8%; Pred. No. 3.7e-08;
Matches 31; Conservative 7; Mismatches 26; Indels 4; Gaps 2;

Qy 50 CYEGNGHYFGKASTDTM--GRCLPWSATVLQOQTYHAHRSDALQGLGKHNCNPON 107
Db 32 CFTANGADYRGTSQNTQTSIDGKPCFLWNE--TFQHPYNTLKYPNGGGLGEHNCNPDG 89

Qy 108 RRRPWCYV 115
Db 90 DVSPWCYV 97

RESULT 28
PLMN_BOVIN
ID PLMN_BOVIN STANDARD; PRT; 812 AA.
AC P06868; Q28162;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7).
GN PLG.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RT Berglund L., Andersen M.D., Petersen T.E.;
RL "Cloning and characterization of the bovine plasminogen cDNA.";
RN Int. Dairy J. 5:593-603(1995).
RX MEDLINE=95203906; PubMed=3845532;
RA Schaller J., Moser P.W., Danneberger-Muller G.A.K., Rosselet S.J., Kampfer U., Rickli E.E.;
RT "Complete amino acid sequence of bovine plasminogen. Comparison with

human plasminogen.";
[3] Eur. J. Biochem. 149:267-278(1985).
SEQUENCE OF 706-812 FROM N.A.
[3] MEDLINE-85023311; PubMed-6148961;
RA Malinowski D.P.; Sadler J.E.; Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RT human and bovine plasminogen."
RL Biochemistry 23:4243-4250(1984).
[4]
RE CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE-88185329; PubMed-3356193;
RA Marti T., Schaller J., Rickli E., Schmid K., Kamerling J.P.,
RA Gerwig G.J., van Halbeek H., Vlieghart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
RT plasminogen. Species specificity in relation to sialylation and
RT fucoylation patterns."
RL Eur. J. Biochem. 173:57-63(1988).
CC -1- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
CC a proteolytic factor in a variety of other processes including
CC embryonic development, tissue remodeling, tumor invasion,
CC and inflammation; in ovulation it weakens the walls of the
CC Graafian follicle. It activates the urokinase-type plasminogen
CC activator, collagenases and several complement zymogens, such
CC as C1 and C5. It cleaves fibrin, fibronectin, thrombospondin,
CC laminin and von Willebrand factor.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -1- ENZYME REGULATION: Converted into plasmin by plasminogen
CC activators, both plasminogen and its activator being bound to
CC fibrin. Cannot be activated with streptokinase.
CC -1- PTM: N-LINKED GLYCANS CONTAIN N-ACETYLACTOSAMINE AND SIALIC ACID.
CC O-LINKED GLYCANS CONSIST OF GAL-GALNAc DISACCHARIDE WITH IS
CC MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).
CC -1- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
CC immediately after dissociation from the clot.
CC -1- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -1- SIMILARITY: Contains 5 kringle domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X79402; CAA65939.1; -;
CC EMBL; K02935; AAA30714.1; -;
CC PIR; S45046; PLEO.
CC HSP; P00747; 2PK4.
CC MEROPS; S01.233; -;
CC GlycoSuiteDB; P06868; -;
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR003014; PAN.
CC InterPro; IPR003609; Pan_app.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC InterPro; IPR003966; Peptidase_S1A_pr.
CC Pfam; PF00051; Kringle; 5.
CC Pfam; PF00024; PAN; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHINOIRIPSPIN.
CC PRINTS; PR00018; KRINGLE.
CC PRINTS; PR01505; PROTHROMBIN.
CC ProDom; PD000395; Kringle; 5.
CC SMART; SM00130; KR; 5.
CC SMART; SM00473; PAN AP; 1.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS00021; KRINGLE_1; 5.
CC PROSITE; PS00070; KRINGLE_2; 5.

DR PROSITE; PS0240; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW Signal.
FT SIGNAL 1 26
FT CHAIN 27 812 PLASMINOGEN.
FT CHAIN 27 583 PLASMIN HEAVY CHAIN A.
FT CHAIN 584 812 PLASMIN LIGHT CHAIN B.
FT DOMAIN 110 188 KRINGLE 1.
FT DOMAIN 192 269 KRINGLE 2.
FT DOMAIN 282 359 KRINGLE 3.
FT DOMAIN 384 461 KRINGLE 4.
FT DOMAIN 485 564 KRINGLE 5.
FT DOMAIN 584 812 SERINE PROTEASE.
FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .).
FT CARBOHYD 365 365 O-LINKED (GALNAc. . .).
FT ACT_SITE 624 624 /FTID-CAR 000015.
FT ACT_SITE 667 667 /FTID-CAR 000015.
FT ACT_SITE 762 762 CHARGE RELAY SYSTEM.
FT ACT_SITE 762 762 CHARGE RELAY SYSTEM.
FT CONFLICT 335 335 N -> D (IN REF. 2).
FT CONFLICT 516 516 Q -> H (IN REF. 2).
FT CONFLICT 555 555 P -> L (IN REF. 2).
FT CONFLICT 744 744 T -> R (IN REF. 3).
SQ SEQUENCE 812 AA; 91216 MW; 38A6A691E220946 CRC64;
Query Match 19.0%; Score 159; DB 1; Length 812;
Best Local Similarity 28.0%; Pred. No. 6.3e-08;
Matches 49; Conservative 15; Mismatches 55; Indels 56; Gaps 11;
QY 5 HQVPSNCDCLN-----GGTCVSNKYPSNIHW--CNCPKKFGQHCIDKSKT 49
Db 315 NRTPEPCKNLEENYCRNPNGEKAFCWYTTN--SEVRWEYCTIPS-----CESSPLST 366
QY 50 -----CYEGNGHFYRGKASTDMGRCPLPMSATVLOQTYHAH---RS 89
Db 367 ERMDVPVPPQTPVQDCYHGNGOSYRGTSSTITGCKQSWSS-----MTPRHLLKTP 421
QY 90 DALQLGLGKNEYCENPNRRPWCYVQGLKPLV--QECMVHDCADGKK--PSSP 140
Db 422 NYFNAGL-TWNYCRNPADKSPWCYT---TDPVRWEFCNLKCKSETPEQVPAAP 472
RESULT 29
ID -ROR1 HUMAN STANDARD; PRT; 937 AA.
AC Q01973; Q92776;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor ROR1 precursor
DE (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 1).
GN ROR1 OR NTRK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX Maslakowski P., Carroll R.D.;
RA Maslakowski P., Carroll R.D.;
RT "A novel family of cell surface receptors with tyrosine kinase-like
RT domain."
RL J. Biol. Chem. 267:26181-26190(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RX Maslakowski P., Carroll R.D.;
RA Maslakowski P., Carroll R.D.;
RT "Human neural tissues express a truncated Ror1 receptor tyrosine
RT kinase, lacking both extracellular and transmembrane domains."

Oncogene 13:1555-1559(1996).

!- FUNCTION: Tyrosine-protein kinase receptor whose role is not yet clear.

!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.

!- SUBCELLULAR LOCATION: Type I membrane protein.

!- ALTERNATIVE PRODUCTS:

Event-Alternative splicing; Named isoforms=2;

Name=Long;

Isoid=Q01973-1; Sequence=Displayed;

Name=Short; Synonym=T-ROR1;

Isoid=Q01973-2; Sequence=VSP_005008;

TISSUE SPECIFICITY: Expressed strongly in human heart, lung, and kidney, but weakly in the CNS. The short isoform is strongly expressed in fetal and adult CNS and in a variety of human cancers, including those originating from CNS or PNS neuroectoderm.

!- DEVELOPMENTAL STAGE: Expressed at high levels during early embryonic development. The expression levels drop strongly around day 16 and there are only very low levels in adult tissues.

!- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR subfamily.

!- SIMILARITY: Contains 1 frizzled (FZ) domain.

!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

!- SIMILARITY: Contains 1 kringle domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

EMBL; M97675; AAC60275.1; -

EMBL; U38894; AAC50714.1; -

PIR; A45082; A45082.

HSP; P00747; ICBA.

Genew; HGNC:10256; ROR1.

MIM; 602335; -

GO; GO:0005737; Cytoplasm; TAS.

GO; GO:0005887; C: integral to plasma membrane; TAS.

GO; GO:0004714; F: transmembrane receptor protein tyrosine kin. . . ; TAS.

GO; GO:0007169; P: transmembrane receptor protein tyrosine kin. . . ; TAS.

InterPro; IPR000024; Fz domain.

InterPro; IPR007110; Ig-like.

InterPro; IPR003598; Ig_c2.

InterPro; IPR000001; Kringle.

InterPro; IPR000719; Prot_kinase.

InterPro; IPR001245; Tyr_kinase.

InterPro; IPR008266; Tyr_kinase_AS.

Pfam; PF01392; Fz; 1.

Pfam; PF00047; ig; 1.

Pfam; PF00051; kringle; 1.

Pfam; PF00699; pkinase; 1.

PRINTS; PR00018; KRINGLE.

PRINTS; PR00109; TYRKINASE.

ProDom; PD000395; Kringle; 1.

ProDom; PD000001; Prot_kinase; 1.

SMART; SM00408; Ig_c2; 1.

SMART; SM00130; KR; 1.

SMART; SM00219; Tyrc; 1.

PROSITE; PS00038; FZ; 1.

PROSITE; PS00835; IG_LIKE; 1.

PROSITE; PS00021; KRINGLE_1; 1.

PROSITE; PS00070; KRINGLE_2; 1.

PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.

PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.

PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

Transferase; Tyrosine-protein kinase; ATP-binding; Receptor; Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation; Immunoglobulin domain; Alternative splicing.

SIGNAL 1 29 POTENTIAL.

FT	CHAIN	30	937	TYROSINE-PROTEIN KINASE TRANSMEMBRANE
FT	RECEPTOR ROR1			
FT	EXTRACELLULAR			
FT	POTENTIAL			
FT	CYTOPLASMIC			
FT	IG-LIKE C2-TYPE			
FT	FZ			
FT	KRINGLE			
FT	PROTEIN KINASE			
FT	SER/THR-RICH			
FT	PRO-RICH			
FT	SER/THR-RICH			
FT	ATP (BY SIMILARITY)			
FT	ATP (BY SIMILARITY)			
FT	ACT SITE			
FT	BY SIMILARITY			
FT	PHOSPHORYLATION (AUTO-)			
FT	BY SIMILARITY			
FT	N-LINKED (GLCNAC)			
FT	N-LINKED (GLCNAC)			
FT	N-LINKED (GLCNAC)			
FT	N-LINKED (GLCNAC)			
FT	Missing (in isoform Short)			
FT	FTId=VSP_005008			
SQ	SEQUENCE	937 AA; 104312 MW; 0D0694DBF29F4773 CRC64;		

Query Match 18.9%; Score 158.5; DB 1; Length 937;

Best Local Similarity 29.7%; Pred. No. 8.1e-08;

Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

QY	2	NELHQP	-----SNCDCLNGTGVSNKYFS-----NIHWNC-----PKKFGGOH 41
Db	236	DETSVVPKRDLCRDECEILENVLCQTEVIFARSNPMLMRKLPNCEDLPQESPEAN	295
QY	42	C-----EIDSKTCYEGNGHYRKGASTDTWGRPCLPWNSATVLOQTYHAHSDAL 92	
Db	296	CIRIGIPMADPINKHKNKYNSTGVDYRGTSVTKSGRQCPWNS-----QYPTHHTFTAL 350	
QY	93	QLG--LGKHNKYNCPDNR--PCWCY 114	
Db	351	RPPELNGCHSYCRNPGNKEAPWCP 375	

RESULT 30

ROR1_MOUSE STANDARD; PRT; 937 AA.

ID_ROR1_MOUSE STANDARD; PRT; 937 AA.

AC Q92139;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Tyrosine-protein kinase transmembrane receptor ROR1 precursor

DE (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 1)

DE (mROR1).

GN ROR1 OR NTRK1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99248426; PubMed=102311392;

RA Oishi I., Takeuchi S., Hashimoto R., Nagabukuro A., Ueda T., Liu Z.J.,

RA Hata T., Akira S., Matsuda Y., Yamamura H., Otani H., Minami Y.;

RT "Spatio-temporally regulated expression of receptor tyrosine kinases,

RT mRor1, mRor2, during mouse development: implications in development

RT and function of the nervous system.";

RL Genes Cells 4:41-56(1999).

CC -!- FUNCTION: Tyrosine-protein kinase receptor whose role is not yet

CC clear.

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein

CC tyrosine phosphate.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR

CC subfamily.

"Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen that forms the linkage between the plasmin chains.";
 Eur. J. Biochem. 58:539-547(1975).
 [9]
 SEQUENCE OF 581-810.
 MEDLINE=77225245; PubMed=142009;
 Kiman B.;
 "Primary structure of the B-chain of human plasmin.";
 Eur. J. Biochem. 76:129-137(1977).
 [10]
 ACTIVE SITE.
 MEDLINE=73149248; PubMed=4694729;
 Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
 "The primary structure of human plasminogen. II. The histidine loop of human plasmin: light (B) chain active center histidine sequence.";
 J. Biol. Chem. 248:1631-1633(1973).
 [11]
 ACTIVE SITE.
 MEDLINE=69234739; PubMed=4240117;
 Groskopf W.R.; Summaria L., Robbins K.C.;
 "Studies on the active center of human plasmin. Partial amino acid sequence of a peptide containing the active center serine residue.";
 J. Biol. Chem. 244:3590-3597(1969).
 [12]
 OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
 MEDLINE=82213905; PubMed=6919539;
 Trexler M., Vali Z., Patthy L.;
 "Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen. Arginine 70 and aspartic acid 56 are essential for binding of ligand by kringle 4.";
 J. Biol. Chem. 257:7401-7406(1982).
 [13]
 FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
 MEDLINE=85054794; PubMed=6094526;
 Vali Z., Patthy L.;
 "The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential for fibrin affinity of the kringle 1 domain.";
 J. Biol. Chem. 259:13690-13694(1984).
 [14]
 PHOSPHORYLATION SITE SER-597.
 MEDLINE=97345939; PubMed=9201958;
 Wang H., Prorok M., Bretthauer R.K., Castellino F.J.;
 "Serine-578 is a major phosphorylation locus in human plasma plasminogen.";
 Biochemistry 36:8100-8106(1997).
 [15]
 CARBOHYDRATE-LINKAGE SITES.
 MEDLINE=88185329; PubMed=3356193;
 Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P., Gerwig G.J., van Halbeek H., Vliegthart J.F.;
 "The N- and O-linked carbohydrate chains of human, bovine and porcine plasminogen. Species specificity in relation to sialylation and fucosylation patterns.";
 Eur. J. Biochem. 173:57-63(1988).
 [16]
 CARBOHYDRATE-LINKAGE SITE SER-268.
 MEDLINE=97207306; PubMed=9054441;
 Pirle-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J., Pizzo S.V.;
 "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of human plasminogen 2.";
 J. Biol. Chem. 272:7408-7411(1997).
 [17]
 CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
 MEDLINE=95042728; PubMed=7525077;
 O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A., Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
 "Angiostatin: a novel angiogenesis inhibitor that mediates the suppression of metastases by a Lewis lung carcinoma.";
 Cell 79:315-328(1994).
 [18]
 CHARACTERIZATION OF ANGIOSTATIN.
 MEDLINE=97238710; PubMed=9102221;

Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W., Lapcevic R., Nacy C.A.;
 "A recombinant human angiotensin protein inhibits experimental primary cancer res.";
 Cancer Res. 57:1329-1334(1997).
 [19]
 X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
 MEDLINE=92031502; PubMed=1657148;
 Mulichak A.M., Tulinsky A., Ravichandran K.G.;
 "Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9-A resolution.";
 Biochemistry 30:10576-10588(1991).
 [20]
 X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
 MEDLINE=92031503; PubMed=1657149;
 Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
 "The refined structure of the epsilon-aminocaproic acid complex of human plasminogen kringle 4.";
 Biochemistry 30:10589-10594(1991).
 [21]
 X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
 Stec B., Yamano A., Whitlow M., Teeter M.M.;
 "Structure of human plasminogen kringle 4 at 1.68 Angstrom and 277 K. A possible structural role of disordered residues.";
 Acta Crystallogr. D 53:169-178(1997).
 [22]
 X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.
 MEDLINE=96180681; PubMed=8611560;
 Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
 "Crystal structures of the recombinant kringle 1 domain of human plasminogen in complexes with the ligands epsilon-aminocaproic acid and trans-4-(aminomethyl)cyclohexane-1-carboxylic acid.";
 Biochemistry 35:2567-2576(1996).
 [23]
 X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
 MEDLINE=98198034; PubMed=9521645;
 Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A., Castellino F.J.;
 "Structure and ligand binding determinants of the recombinant kringle 5 domain of human plasminogen.";
 Biochemistry 37:3258-3271(1998).
 [24]
 STRUCTURE BY NMR OF 96-184.
 MEDLINE=94237157; PubMed=8181475;
 Rejante M.R., Llinas M.;
 "1H-NMR assignments and secondary structure of human plasminogen kringle 1.";
 Eur. J. Biochem. 221:927-937(1994).
 [25]
 STRUCTURE BY NMR OF 96-184.
 MEDLINE=94237158; PubMed=8181476;
 Rejante M.R., Llinas M.;
 "Solution structure of the epsilon-aminohexanoic acid complex of human plasminogen kringle 1.";
 Eur. J. Biochem. 221:939-949(1994).
 [26]
 STRUCTURE BY NMR OF 183-354.
 MEDLINE=96194156; PubMed=8652577;
 Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M., Rickli E.E.;
 "Recombinant gene expression and 1H NMR characteristics of the kringle (2 + 3) supermodule: spectroscopic/functional individuality of plasminogen kringle domains.";
 Biochemistry 35:2357-2364(1996).
 [27]
 STRUCTURE BY NMR OF 374-461.
 MEDLINE=90219023; PubMed=2157850;
 Query Match 18.7%; Score 156.5; DB 1; Length 810;
 Best Local Similarity 28.5%; Pred. NO. 1.1e-07;
 Matches 49; Conservative 22; Mismatches 52; Indels 49; Gaps 12;
 QY 5 HQVPSNCDCLNGTGVSNKYFSN-----IHWNCNP-KKFGQHCET---DKS----- 47

308 NRTPEFPCKN-----LDENYCNPDGKRAPWCHTNSQVRYEYCKIPSCDSFVSTEQIA 363
48 -----KTCYGNHGFRGKASTMTGRPCLPNSATVLOQTVAHR-----SDALQ 93
364 PTAPPETLPVQDCYRGDSQVAGTSTTTGKCKQSWSS-----MTRHOKTPTENYFN 418
94 LGIGKKNYCNPNRRPWCYQVGLKPLV--QECNVHDCADGKPS--SPP 141
419 AGL-TWNYCNPDADGKPCFT---TDPVRYEYCNLKKCS-GTASVAVPP 465

RE [T 32
KX HUMAN
ID KRM1 HUMAN STANDARD; PRT; 475 AA.
AC Q96MJB; Q96V70; Q9UGS5; Q9UGU1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (kringle-containing protein marking the eye
and the nose) (Dickkopf receptor).
GN KREMN1 OR KREMN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RF SEQUENCE FROM N.A. (ISOFORM 1).
RA Nakamura T., Nakamura T.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RF SEQUENCE FROM N.A. (ISOFORM 2).
RA Tashiro H., Yamazaki M., Watanabe K., Kunagai A., Itakura S.,
RA Otsukumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Negai K.,
RA Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RF SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Begguley C., Bailey J.E., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carter C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Copley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcott R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA McLaren I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spraggon L., Stewart C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoeshima S., Kawaasaki K., Sasaki T., Asakawa S., Kudo H.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyma S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hua A., Kerton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,

Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
Zhan M., Zhang G., Chisoso S., Murray J., Miller N., Minx P.,
Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
Hinds K., Kamp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
Scheet P., Walker C., Wansley A., Wohldmann P., Pepin K., Nelson J.,
Korfi I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
Budarf M.L., Mcdermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
Peyraud M., Kieda D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
Khan A.S., Lane L., Tilahun Y., Wright H.,
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
CC -!- FUNCTION: receptor for Dickkopf protein. Cooperates with Dickkopf
to block Wnt/beta-catenin signaling (By similarity). (Potential).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q96MJB-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q96MJB-2; Sequence=VSP_003900;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB059618; BAB40369.1; -;
DR EMBL; AK056425; BAB71180.1; -;
DR EMBL; Z95116; CAB62952.1; -;
DR EMBL; AL021393; CAB62959.1; -;
DR GO; HGNC:17550; KREMN1.
DR GO; GO:0016021; C:integral to membrane; ISS.
DR GO; GO:0005624; C:membrane fraction; TAS.
DR GO; GO:0007154; P:cell communication; TAS.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR000003; Kringle.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
KW Wnt signaling pathway; Signal; Transmembrane; Kringle;
KW Alternative splicing;
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 475 KREMN PROTEIN 1.
FT DOMAIN 21 394 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 395 415 POTENTIAL.
FT DOMAIN 416 475 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 33 116 KRINGLE.
FT DOMAIN 121 202 WSC.
FT DOMAIN 216 323 CUB.
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).

10-OCT-2003 (Rel. 42, Last sequence update)
15-MAR-2004 (Rel. 43, Last annotation update)
Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
PLG.
Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=91184812; PubMed=2081600;
RA Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;
RT "Characterization of the cDNA coding for mouse plasminogen and
RT localization of the gene to mouse chromosome 17.";
RL Genomics 8:49-61(1990).
[2]
SEQUENCE FROM N.A.
RX STRAIN=129/Sv;
RA Brathwaite M., Waeltz P., Qian Y., Dudekula D., Schlessinger D.,
RA Nagaraja R.;
RT "Genomic sequence analysis in the mouse t-complex region.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RX TISSUE=Liver;
RA MEDLINE=22388257; PubMed=12477932;
RA Strausberg K.L., Feilgold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marudana K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
SEQUENCE OF 1-16 FROM N.A.
RX STRAIN=129/SvJ; TISSUE=Liver;
RA Barnach F.G., Gutierrez A., Fowler B.J., Bugge T.H., Degen J.L.,
RA Farmer R.J., Miles L.A.;
RT "Localization of regulatory elements mediating constitutive and
RT cytokine-stimulated plasminogen gene expression.";
RL J. Biol. Chem. 277:38579-38588(2002).
[5]
CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328(1994).
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
CC a proteolytic factor in a variety of other processes including
CC embryonic development, tissue remodeling, tumor invasion, and
CC inflammatory; in ovulation it weakens the walls of the Graafian
CC follicle. It activates the urokinase-type plasminogen activator,
CC collagenases and several complement zymogens, such as C1 and C5.
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von
CC Willebrand factor.
CC -!- FUNCTION: Angiostatin is an angiogenesis inhibitor that blocks
CC neovascularization and growth of experimental primary and

metastatic tumors in vivo.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen
CC activators, both plasminogen and its activator being bound to
CC fibrin. Cannot be activated with streptokinase.
CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
CC immediately after dissociation from the clot.
CC -!- MISCELLANEOUS: In the presence of the inhibitor, the activation
CC involves only cleavage after Arg-581, resulting in 2 chains held
CC together by 2 disulfide bonds. Without the inhibitor, the
CC activation involves also removal of the activation peptide.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -!- SIMILARITY: Contains 5 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J04766; AA50168.1; -
CC EMBL; AF481053; AM22156.1; -
CC EMBL; BC014773; AAH14773.1; -
CC EMBL; BC057186; AAH57186.1; -
CC EMBL; AY134430; AAH15805.1; -
CC PIR; A38514; PLMS.
CC HSP; P00747; IPMK.
CC MEROPS; S01.233; -
CC MGD; MGI:97620; Plg.
CC GO; GO:0016506; Fibrinolysis activator activity; IDA.
CC GO; GO:0006915; Apoptosis; IDA.
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR003014; PAN.
CC InterPro; IPR003609; Pan app.
CC InterPro; IPR001254; Peptidase S1.
CC InterPro; IPR001314; Peptidase S1A.
CC InterPro; IPR003966; Peptidase_S1A_pr.
CC Pfam; PF00051; kringle; 5.
CC Pfam; PF00024; PAN; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC PRINTS; PR01505; PROTHROMBIN.
CC ProDom; PD000395; Kringle; 4.
CC SMART; SM00130; KR; 5.
CC SMART; SM00473; PAN AP; 1.
CC SMART; SM00020; Tryp Spc; 1.
CC PROSITE; PS00021; KRINGLE_1; 5.
CC PROSITE; PS00070; KRINGLE_2; 5.
CC PROSITE; PS02400; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC KW Hydrolyase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
CC Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
CC Signal.
CC FT CHAIN 1 19 PLASMINOGEN.
CC FT CHAIN 20 812 PLASMIN HEAVY CHAIN A.
CC FT PEPTIDE 20 97 ACTIVATION PEPTIDE.
CC FT CHAIN 98 581 PLASMIN SHORT FORM OF CHAIN A.
CC FT CHAIN 98 7436 ANGIOSTATIN.
CC FT CHAIN 582 812 PLASMIN LIGHT CHAIN B.
CC FT DOMAIN 103 181 KRINGLE 1.
CC FT DOMAIN 184 262 KRINGLE 2.
CC FT DOMAIN 275 352 KRINGLE 3.
CC FT DOMAIN 377 454 KRINGLE 4.
CC FT DOMAIN 481 560 KRINGLE 5.
CC FT DOMAIN 582 812 SERINE PROTEASE.

ACT SITE 624 CHARGE RELAY SYSTEM (BY SIMILARITY).
 ACT SITE 667 CHARGE RELAY SYSTEM (BY SIMILARITY).
 ACT SITE 762 CHARGE RELAY SYSTEM (BY SIMILARITY).
 DISULFID 49 BY SIMILARITY.
 DISULFID 53 BY SIMILARITY.
 DISULFID 103 BY SIMILARITY.
 DISULFID 124 BY SIMILARITY.
 DISULFID 152 BY SIMILARITY.
 DISULFID 185 BY SIMILARITY.
 DISULFID 188 BY SIMILARITY.
 DISULFID 206 BY SIMILARITY.
 DISULFID 234 BY SIMILARITY.
 DISULFID 275 BY SIMILARITY.
 DISULFID 326 BY SIMILARITY.
 DISULFID 344 BY SIMILARITY.
 DISULFID 377 BY SIMILARITY.
 DISULFID 398 BY SIMILARITY.
 DISULFID 426 BY SIMILARITY.
 DISULFID 481 BY SIMILARITY.
 DISULFID 502 BY SIMILARITY.
 DISULFID 531 BY SIMILARITY.
 DISULFID 568 INTERCHAIN (BY SIMILARITY).
 DISULFID 578 INTERCHAIN (BY SIMILARITY).
 DISULFID 609 BY SIMILARITY.
 DISULFID 701 BY SIMILARITY.
 DISULFID 731 BY SIMILARITY.
 DISULFID 758 BY SIMILARITY.
 CONFLICT 235 R -> H (IN REF. 1).
 CONFLICT 525 S -> D (IN REF. 1).
 CONFLICT 649 S -> L (IN REF. 1).
 SEQUENCE 812 AA; 90781 MW; 2417326086A2FFD2 CRC64;

18.5%; Score 155; DB 1; Length 812;
 est Local Similarity 30.1%; Pred. No. 1.5e-07;
 atches 40; Conservative 12; Mismatches 55; Indels 26; Gaps 7;

30 WC--NCPKFGGHCHEIDKSKT-----CYEGNGHFYRGKASTDTMGRPCLPWNSA 77
 244 WCFTTDPK-RWEYCDIPRCTTPPPPTQYCLKRGYKRVTSVTSVSGKTQCRWS-- 300
 78 TVLQCTYHARDALQLGLG--KHNYCRNPDRRPWCYVQGLKPLVQECWHDGADGK 135
 301 ----EOTPHRHNTFNPCKNLEENYCRNPDGETAPWCYT-TDSQLRWEYCEIFPCESSA 356
 136 KP-----SSPPE 143
 357 SPDQSDSSVPPPE 369

PLN ERIEU STANDARD; PRT; 810 AA.
 01-NOV-1997 (Rel. 35, Created)
 01-NOV-1997 (Rel. 35, Last sequence update)
 28-FEB-2003 (Rel. 41, Last annotation update)
 Plasminogen precursor (EC 3.4.21.7).
 PLG.
 O Brinaceus europaeus (Western European hedgehog).
 O Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 O Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceinae; Erinaceus.
 O NCBI_TaxID=9345;
 [1]
 SEQUENCE FROM N.A.
 TISSUE=Liver;
 MEDLINE=96025778; PubMed=7592597;
 Lawn R.M., Boonmark N.W., Schwartz K., Lindahl G.E., Wade D.P.,
 Byrne C.D., Fong K.J., Meer K., Pathy L.;
 "The recurring evolution of lipoprotein(a). Insights from cloning of
 hedgehog apolipoprotein(a).";
 J. Biol. Chem. 270:24004-24009 (1995).
 [2]
 REVISIONS.

Lawn R.M.;
 Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 -1- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
 a proteolytic factor in a variety of other processes including
 embryonic development, tissue remodeling, tumor invasion, and
 inflammation; in ovulation it weakens the walls of the Graafian
 follicle. It activates the urokinase-type plasminogen activator,
 collagenases and several complement zymogens, such as C1 and C5.
 It cleaves fibrin, fibronectin, thrombospondin, laminin and von
 Willebrand factor.
 -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 higher selectivity than trypsin. Converts fibrin into soluble
 products.
 -1- ENZYME REGULATION: Converted into plasmin by plasminogen
 activators, both plasminogen and its activator being bound to
 fibrin. Cannot be activated with streptokinase.
 -1- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
 immediately after dissociation from the clot.
 -1- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
 -1- SIMILARITY: Contains 5 kringle domains.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; U33171; AAC48717.1; -.
 PIR; I46260; I46260.
 HSP; P00747; IPMK.
 MROPS; S01.233.
 InterPro: IPR009003; Cys_Ser_trypsin.
 InterPro: IPR000001; Kringle.
 InterPro: IPR003014; PAN.
 InterPro: IPR003609; Pan app.
 InterPro: IPR012514; Peptidase S1.
 InterPro: IPR001314; Peptidase_S1A.
 InterPro: IPR003966; Peptidase_S1A_pr.
 Pfam; PF00051; Kringle; 5.
 Pfam; PF00024; PAN; 1.
 Pfam; PF00089; trypsin; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 PRINTS; PR00018; KRINGLE.
 PRINTS; PR01505; PROTHROMBIN.
 ProDom; PD000395; Kringle; 5.
 SMART; SMC0130; KR; 5.
 SMART; SMC0473; PAN AP; 1.
 SMART; SMC0020; Tryp_Spc; 1.
 PROSITE; PS00021; KRINGLE_1; 5.
 PROSITE; PS00070; KRINGLE_2; 5.
 PROSITE; PS00240; TRYPsin_DOM; 1.
 PROSITE; PS00134; TRYPsin_HIS; 1.
 PROSITE; PS00135; TRYPsin_SER; 1.
 KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
 KW Signal.
 FT SIGNAL 1 19 BY SIMILARITY.
 FT CHAIN 20 810 PLASMINOGEN.
 FT CHAIN 20 582 PLASMIN HEAVY CHAIN A (BY SIMILARITY).
 FT CHAIN 583 810 PLASMIN LIGHT CHAIN B (BY SIMILARITY).
 FT DOMAIN 103 181 SERINE PROTEASE.
 FT DOMAIN 185 262 KRINGLE 1.
 FT DOMAIN 275 352 KRINGLE 2.
 FT DOMAIN 379 456 KRINGLE 3.
 FT DOMAIN 482 561 KRINGLE 4.
 FT ACT_SITE 622 622 CHARGE RELAY SYSTEM.
 FT ACT_SITE 665 665 CHARGE RELAY SYSTEM.
 FT ACT_SITE 760 760 CHARGE RELAY SYSTEM.
 FT CARBOHYD 339 339 N-LINKED (GLCNAC...) (POTENTIAL).
 SEQUENCE 810 AA; 90902 MW; 8E75780946017A16 CRC64;

Query Match 18.3%; Score 153.5; DB 1; Length 810;
 Best Local Similarity 27.3%; Pred. No. 2.1e-07;
 Matches 39; Conservative 12; Mismatches 47; Indels 45; Gaps 6;

QY 6 QVPSNCCILN-----GTCVSNKYFSNIHW--CNCPKKFGQCHCEIDKSKT- 49
 Db 309 RTPENYPCXNLDENYCRNPDGPAPWCFTTN--SSVRWEFCXIP-----DCVSSASETE 360
 QY 50 -----CYEGNGHFYRCKASTDTMGPRCLPWNASTVLQTYHAHRSDA 91
 Db 361 HSDAPVIVPEPTFVVQECYQNGQTYRGTSTTTTGKCKQCPWTSRMRPHRSHKTPENYPD 420

QY 92 LQGLGKKNYCRNPDNRPPWCY 114
 Db 421 ADLTM--NYCRNPDGKGPWCY 440

RE UT 36
 FLA RAT
 ID PLMN RAT STANDARD; PRT; 169 AA.
 AC Q01177;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasminogen (EC 3.4.21.7) (Fragment).
 GN PLG.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=91250378; PubMed=1645711;
 RA Kanakas J.J., Makker S.P.;
 RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
 RL J. Biol. Chem. 266:10825-10829 (1991).
 CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
 CC a proteolytic factor in a variety of other processes including
 CC embryonic development, tissue remodeling, tumor invasion, and
 CC inflammation; in ovulation it weakens the walls of the Graafian
 CC follicle. It activates the urokinase-type plasminogen activator,
 CC collagenases and several complement zymogens, such as C1 and C5.
 CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von
 CC Willebrand factor.
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen
 CC activators, both plasminogen and its activator being bound to
 CC fibrin. Cannot be activated with streptokinase.
 CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
 CC immediately after dissociation from the clot.
 CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
 CC -!- SIMILARITY: Contains 5 kringle domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; M62832; AAA41884.1; -;
 PIR; A40522; A40522.
 HSSP; P00747; IPMK.
 MEROPS; S01.233; -;
 InterPro; IPR000001; Kringle.
 InterPro; IPR001254; Peptidase_S1.
 InterPro; IPR003966; Peptidase_S1A_pr.

DR Pfam: PF00051; kringle; 2.
 DR PRINTS; PRO0018; KRINGLE.
 DR PRINTS; PRO1505; PROTHROMBIN.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 2.
 DR PROSITE; PS00240; TRYPsin_DOM; PARTIAL.
 DR PROSITE; PS00134; TRYPsin_HIS; PARTIAL.
 DR PROSITE; PS00135; TRYPsin_SER; PARTIAL.
 KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Repeat.
 FT NON TER 1
 FT DOMAIN <1 10 KRINGLE 3 (BY SIMILARITY).
 FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
 FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).
 FT DISULFID 34 112 BY SIMILARITY.
 FT DISULFID 55 95 BY SIMILARITY.
 FT DISULFID 83 107 BY SIMILARITY.
 FT NON TER 169 169
 SQ SEQUENCE 169 AA; 18401 MW; 77A54214C49D010C CRC64;

Query Match 18.2%; Score 152.5; DB 1; Length 169;
 Best Local Similarity 30.2%; Pred. No. 5.6e-08;
 Matches 42; Conservative 18; Mismatches 32; Indels 47; Gaps 9;

QY 40 QHCEI-----DKS-----KTCYEGNGHFYRCKASTDTMGPRCLPWNAST 78
 Db 3 EYCEIPSCGSSVPDQSDSVLPETQPVVQECYQNGKSYRGTSTTTNGKQSW--- 58
 QY 79 VLQQTVAHESDALQL---GLGKKNYCRNPDN--RRPWCYVQVGLKPLV--QECMVHDC 132
 Db 59 -VSMTPHSHSKTTPANFPDGL-EMNYCRNPDNDQRGWCFT---TDPVSRWEYCNLRCS 113

QY 133 D-----GKFPSSP 140
 Db 114 ETGGGVAESAIVPQVPSAP 132

RESULT 37
 HGFL HUMAN STANDARD; PRT; 711 AA.
 ID HGFL HUMAN
 AC P26927; Q13350; Q14870;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hepatocyte growth factor-like protein precursor (Macrophage
 DE stimulatory protein) (MSP) (Macrophage stimulating protein).
 GN MST1 OR HGFL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=92002016; PubMed=1655021;
 RA Han S., Stuart L.A., Friesner Degen S.J.;
 RT "Characterization of the DNP152 locus on human chromosome 3:
 RT identification of a gene coding for four kringle domains with
 RT homology to hepatocyte growth factor.";
 RL Biochemistry 30:9768-9780 (1991).
 CC (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93340141; PubMed=8393443;
 RA Yoshimura T., Yuhki N., Wang M.H., Skeel A., Leonard E.J.;
 RT "Cloning, sequencing, and expression of human macrophage stimulating
 RT protein (MSP, MST1) confirms MSP as a member of the family of kringle
 RL J. Biol. Chem. 268:15461-15468 (1993).
 CC -!- FUNCTION: Probably has no proteolytic activity, since crucial AA
 CC characteristic of serine proteases catalytic sites are not


```

FT CARBOHYD 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 13 Y -> C.
FT 212 /FTID=VAR_006631.
FT 212 C -> F.
FT 676 /FTID=VAR_006632.
FT 676 E -> K (in dbSNP:7798).
FT 623 /FTID=VAR_014569.
FT CONFLICT 623 L -> F (IN REF. 2).
FT SEQUENCE 711 AA; 80379 MW; 596ED21F180290EA CRC64;

Query Match 17.9%; Score 150; DB 1; Length 711;
Best Local Similarity 27.0%; Pred. No. 4.1e-07;
Matches 34; Conservative 20; Mismatches 46; Indels 26; Gaps 6;

QY 30 WC-NCPKKGQGHCEIDK-----SKTYEGNGHFIYFKASTDTMGKPCLPWNS 76
Db 250 WCYTDPIQIERFCDLFRGSEAPQREATTVSCFRKGEGYGTANTTAGVPCQRWDA 309
QY 77 ATVLQQTYYAHRSDALQLGLG--KHYNCNPDRNRFPWCYVQVGLKPLVQEC---MVHDC 131
Db 310 ----QIPHQRFTPEKYACKDLRFNCFNPDGSEAPWCFT---LRPGMEAFYQIYRRC 361
QY 132 ADGKKP 137
Db 362 TDDVRP 367

RESULT 38
KRM2 HUMAN
ID _KRM2_HUMAN STANDARD; PRT: 462 AA.
AC Q8NCW0; Q8NCJ4; Q8NCW1; Q96GL8; Q9BTP9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kremen protein 2 precursor [Kringele-containing protein marking the eye
DE and the nose] (Dickkopf receptor 2).
OS KREMEN2 OR KRM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RA Tanaka S., Sugimachi K.;
RT "Human Kremen2 and Wnt signaling."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RA Tissue/Ovarian Carcinoma;
RC Isegai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwavanagi T., Ninomiya K.;
RT "NEDD human cDNA sequencing project."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RC Tissue=Brain, and Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwain K.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Rahav J., Hailton E., Khatami M., Madan A.,

```

```

Whiting M., Madan A., Young A.C., Shervchenko Y., Bouffard G.G.,
Blakesley R.W., Grimwood J.W., Green E.D., Dickinson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.:
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
-!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
to block Wnt/beta-catenin signaling. Forms a ternary complex with
Dkk1 and LRP6 and induces rapid endocytosis and removal of the Wnt
receptor LRP6 from the plasma membrane (By similarity).
-!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=4;
Name=1;
IsoId=Q8NCW0-1; Sequence=Displayed;
Name=2; Synonyms=Kremen2a;
IsoId=Q8NCW0-2; Sequence=VSP_050509, VSP_050510;
Name=3; Synonyms=Kremen2b;
IsoId=Q8NCW0-3; Sequence=VSP_050511, VSP_050512;
Name=4; Synonyms=Kremen2c;
IsoId=Q8NCW0-4; Sequence=VSP_050513, VSP_050514;
-!- SIMILARITY: Contains 1 CUB domain.
-!- SIMILARITY: Contains 1 kringle domain.
-!- SIMILARITY: Contains 1 WSC domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----

```

[illegible]

C -1- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
 C -1- SIMILARITY: Contains at least 1 kringle domain.
 D HSPF, F00747; SHFG.
 D MEROPS, S01233; --
 D InterPro; IPR009003; Cys_Ser_tryptase.
 D InterPro; IPR000001; Kringle.
 D InterPro; IPR001254; Peptidase S1.
 D InterPro; IPR001314; Peptidase_SIA.
 D InterPro; IPR003966; Peptidase_SIA_pr.
 D Pfam; PF00051; kringle; 1.
 D Pfam; PF00089; trypsin; 1.
 D PRINTS; PR00722; CHYNOTRYPIN.
 D PRINTS; PR00018; KRINGLE.
 D PRINTS; PR00134; TRYPSIN_DOM; 1.
 D PRINTS; PR01505; PROTHROMBIN.
 D ProDom; PD000395; Kringle; 1.
 D SMART; SM00130; KR; 1.
 D SMART; SM00026; Tryp_SPc; 1.
 D PROSITE; PS00021; KRINGLE_1; 1.
 D PROSITE; PS00070; KRINGLE_2; 1.
 D PROSITE; PS00240; TRYPSIN_DOM; 1.
 D PROSITE; PS00134; TRYPSIN_HIS; 1.
 D PROSITE; PS00135; TRYPSIN_SER; 1.
 K Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 K Tissue remodeling; blood coagulation; Kringle; Zymogen.
 F NON_TER 1
 F CHAIN <1 103 PLASMIN HEAVY CHAIN A.
 F CHAIN 104 333 PLASMIN LIGHT CHAIN B.
 F DOMAIN 104 83 KRINGLE 5.
 F DOMAIN 104 333 SERINE PROTEASE.
 F DISULFID 4 83 BY SIMILARITY.
 F DISULFID 25 66 BY SIMILARITY.
 F DISULFID 34 78 BY SIMILARITY.
 F DISULFID 90 208 INTERCHAIN (BY SIMILARITY).
 F DISULFID 100 108 INTERCHAIN (BY SIMILARITY).
 F DISULFID 130 146 BY SIMILARITY.
 F DISULFID 222 289 BY SIMILARITY.
 F DISULFID 252 268 BY SIMILARITY.
 F DISULFID 279 307 BY SIMILARITY.
 F ACT_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).
 F ACT_SITE 188 188 CHARGE RELAY SYSTEM (BY SIMILARITY).
 F ACT_SITE 283 283 CHARGE RELAY SYSTEM (BY SIMILARITY).
 F SITE 152 152 STREPTOKINASE-BINDING SITE (PROBABLE).
 F SITE 186 186 STREPTOKINASE-BINDING SITE (PROBABLE).
 F SITE 264 264 STREPTOKINASE-BINDING SITE (PROBABLE).
 F SITE 277 277 SITE OF SUBSTRATE SPECIFICITY (BY SIMILARITY).
 S SEQUENCE 333 AA; 36678 MW; C8C0271B6C6AC8D4 CRC64;
 Very Match 17.7%; Score 148; DB 1; Length 333;
 est Local Similarity 38.8%; Pred. No. 3e-07;
 atches 40; Conservative 7; Mismatches 40; Indels 16; Gaps 6;
 Q 50 CYEGNGHFVGRKASTDTWGRPCLPWNSATVLOQTHAHR----SDALQGLGKKNYCRNP 105
 D 4 CMFNGKGYRGKATVNGIPQEWAA-----QEPHRSIFTPETNPQAGLEK-NYCRNP 57
 Q 106 D-NRRRPWCYOVGLKPLVQECMWHDC-----DGKKPSPPPE 143
 D 58 DDVNGPWCYT-MNQKLFYDCVPCQVSTSFDCGKQVPEKK 99
 R JLT 40
 R 2 HUMAN
 I ROR2 HUMAN STANDARD; PRT; 943 AA.
 A Q01974; Q9HAY7; Q9HB61;
 D 16-OCT-2001 (Rel. 40, Created)
 D 16-OCT-2001 (Rel. 40, Last sequence update)
 D 15-MAR-2004 (Rel. 43, Last annotation update)
 D Tyrosine-protein kinase transmembrane receptor ROR2 precursor
 D (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 2).
 G ROR2 OR NTRK2.
 C Homo sapiens (Human).
 C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 O

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93100347; PubMed=1334494;
 RA Maslakowski P., Carroll R.D.;
 RT "A novel family of cell surface receptors with tyrosine kinase-like domain.";
 RL J. Biol. Chem. 267:26181-26190(1992).
 RN [2]
 RP SEQUENCE OF 34-943 FROM N.A., AND VARIANT ILE-819.
 RX MEDLINE=20164326; PubMed=10700182;
 RA Oldridge M., Fortuna A.M., Maringa M., Propping P., Mansour S.,
 RA Pollitt C., DeChiara T.M., Kimble R.B., Valenzuela D.M.,
 RA Yancopoulos G.D., Wilkie A.O.M.;
 RT "Dominant mutations in ROR2, encoding an orphan receptor tyrosine kinase, cause brachydactyly type B.";
 RL Nat. Genet. 24:275-278(2000).
 RN [3]
 RP SEQUENCE OF 34-574 FROM N.A., AND VARIANT THR-245.
 RX MEDLINE=20442029; PubMed=10986040;
 RA Schwabe G.C., Tinschert S., Buschow C., Meinecke P., Wolff G.,
 RA Gillespie-Kaeschbach G., Oldridge M., Wilkie A.O.M., Koemec R.,
 RA Mundlos S.;
 RT "Distinct mutations in the receptor tyrosine kinase gene ROR2 cause brachydactyly type B.";
 RL Am. J. Hum. Genet. 67:822-831(2000).
 RN [4]
 RP VARIANTS RRS CYS-184; TRP-189; TRP-366 AND LYS-620.
 RX MEDLINE=20392394; PubMed=10932186;
 RA Azzi A.R., Rajad A., Fenske C.D., Oldridge M., Elanko N.,
 RA Teres-Ferreira E., Tuysuz B., Murday V.A., Patton M.A.,
 RA Wilkie A.O.M., Jeffery S.;
 RT "Recessive Robinow syndrome, allelic to dominant brachydactyly type B, is caused by mutation of ROR2.";
 RL Nat. Genet. 25:419-422(2000).
 RN [5]
 RP VARIANT RRS TYR-182.
 RX MEDLINE=20392395; PubMed=10932187;
 RA van Bokhoven H., Celli J., Kayserili H., van Beusekom E., Balci S.,
 RA Brussel W., Skovby F., Kerr B., Percin E.F., Akarsu N., Brunner H.G.;
 RL Nat. Genet. 26:383-383(2000).
 CC -1- FUNCTION: Tyrosine-protein kinase receptor which may be involved in the early formation of the chondrocytes. It seems to be required for cartilage and growth plate development.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- DEVELOPMENTAL STAGE: Expressed at high levels during early embryonic development. The expression levels drop strongly around day 16 and there are only very low levels in adult tissues.
 CC -1- DISEASE: Defects in ROR2 are a cause of brachydactyly type B1 (BDB1) [MIM:113000]. BDB1 is an autosomal dominant skeletal disorder characterized by hypoplasia/aplasia of distal phalanges and nails. In BDB1 the middle phalanges are short but in addition the terminal phalanges are rudimentary or absent. Both fingers and toes are affected. The thumbs and big toes are usually deformed.
 CC -1- DISEASE: Defects in ROR2 are a cause of recessive Robinow syndrome (RRS) [MIM:26810]. RRS is an autosomal disorder characterized by skeletal dysplasia with generalized limb bone shortening, segmental defects of the spine, brachydactyly and a dysmorphic facial appearance.
 CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR subfamily.
 CC -1- SIMILARITY: Contains 1 frizzled (FZ) domain.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.


```

FT CARBOHYD 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 944 AA; 105050 MW; CD2EEBC710387A56 CRC64;

Query Match 17.7%; Score 148; DB 1; Length 944;
Best Local Similarity 36.0%; Pred. No. 8 6e-07;
Matches 40; Conservative 11; Mismatches 36; Indels 24; Gaps 8;

QY 32 NC-----PKKFGGQHCIEDKSKTCYEGNGHFYRGKASTDTMGRPCLPNNSATVLCQTYH 85
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 NCMRIGIPAERLGRYH-----QCYNAGADYRGVASTTKSHQCQPW---ALQHP-H 346
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 86 AHR---SDALQGLGKHNYCRPNDRER--PWCYVVOVGLKRLVQECVHDC A 132
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 347 SHRLSTPEPLG-GGHAYCRNPGGQMEGFWCPTQ-NKNVRVELCDVPPCS 395
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 42
HGFL MOUSE STANDARD; PRT; 716 AA.
AC P26928;
AT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hepatocyte growth factor-like protein precursor (Macrophage
DE stimulatory protein) (MSP).
GN MST1 OR HGFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/c; TISSUE=Liver;
RX MEDLINE=92002017; PubMed=1832957;
RA Friesner Degen S.J., Stuart L.A., Han S., Jamison C.S.;
RT "Characterization of the mouse cDNA and gene coding for a hepatocyte
RT growth factor-like protein: expression during development.";
RL Biochemistry 30:9781-9791(1991).
CC -1- FUNCTION: Probably has no proteolytic activity, since crucial AA
CC characteristic of serine proteases catalytic sites are not
CC conserved.
CC -1- TISSUE SPECIFICITY: Liver. Lower levels in lung, placenta and
CC adrenal.
CC -1- DEVELOPMENTAL STAGE: Is expressed at low levels during gestation.
CC Just before birth the level increases dramatically and remains
CC stable afterwards.
CC -1- PTM: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE
CC HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE
CC POLYPEPTIDES.
CC -1- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -1- SIMILARITY: Contains 4 kringle domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M74180; AAA50166.1; -
CC EMBL; M74181; AAA50167.1; -
CC PIR; A40332; A40332.
CC HSP; P00747; IKNR.
CC MEROPS; S01.975; -.
CC MGD; MGI.96080; Mst1.
CC GO; GO:0007566; Embryo implantation; IC.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR003014; PAN.
CC InterPro; IPR003609; Pan_app.
CC InterPro; IPR001254; Pentidase S1.

```

DR InterPro: IPR001314; Peptidase S1A.
 DR InterPro: IPR003966; Peptidase S1A_pr.
 DR Pfam: PF00051; Kringle; 4.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR PRINTS: PR01505; PROTHROMBIN.
 DR ProDom: PD000395; Kringle; 4.
 DR SMART: SM00130; KR; 4.
 DR SMART: SM00473; PAN_APC; 1.
 DR SMART: SM00020; TRYD_SPC; 1.
 DR PROSITE: PS00021; KRINGLE 1; 4.
 DR PROSITE: PS00070; KRINGLE 2; 4.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR Kringle; Glycoprotein; Serine protease homolog; Repeat; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 716
 FT DOMAIN 19 109
 FT DOMAIN 110 186
 FT DOMAIN 111 268
 FT DOMAIN 192 370
 FT DOMAIN 379 457
 FT DOMAIN 489 716
 FT DISULFID 56 78
 FT DISULFID 60 66
 FT DISULFID 110 186
 FT DISULFID 131 169
 FT DISULFID 157 181
 FT DISULFID 191 268
 FT DISULFID 194 333
 FT DISULFID 212 251
 FT DISULFID 240 263
 FT DISULFID 292 370
 FT DISULFID 313 352
 FT DISULFID 341 364
 FT DISULFID 379 457
 FT DISULFID 400 440
 FT DISULFID 428 452
 FT DISULFID 477 593
 FT DISULFID 512 528
 FT DISULFID 607 672
 FT DISULFID 637 651
 FT DISULFID 662 690
 FT CARBOHYD 72 72
 FT CARBOHYD 173 173
 FT CARBOHYD 305 305
 FT CARBOHYD 620 620
 FT CONFLICT 19 19
 FT SEQUENCE 716 AA; 80588 MW; BICE02EF85213ACC CRC64;
 Query Match 17.6%; Score 147.5; DB 1; Length 716;
 Best Local Similarity 36.5%; Pred. No. 7.3e-07;
 Matches 38; Conservative 11; Mismatches 34; Indels 21; Gaps 7;
 QY 50 CYEGNGHFYRGKASTDITMGRPCLPNSATVLOQTYHAHR-----SDALQLGLGKHNYCRN 104
 DB 379 CYHSGEQYRGVSKTRGVQCHWSSET-----PHKQFTPTSPAQGL-EANFCRN 430
 QY 105 PD-NRRPWCYVQVGLKP--LYQECWHDCAKCKPS--SPPEE 143
 DB 431 PDGSHGHPWCYT---LDPDILFDYCALORCDDDPSPSLDPPDQ 471
 RE: 43
 KR: MOUSE
 ID: KR2 MOUSE STANDARD; PRT; 461 AA.
 AC: O8K157;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Kremen protein 2 precursor (Kringle-containing protein marking the eye
 and the nose) (Dickkopf receptor 2).

GN KREMEN2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND FUNCTION.
 RC STRAIN=CS7BL/6J;
 RX MEDLINE=22045977; PubMed=12050670;
 RA Mac B., Wu W., Davidson G., Marhold J., Li M., Mechler B.M.,
 RA Delius H., Hoppe D., Stannek P., Walter C., Glinka A., Niehrs C.;
 RT "Kremen proteins are Dickkopf receptors that regulate Wnt/beta-catenin
 signalling";
 EL Nature 417:664-667(2002).
 CC -!- FUNCTION: Receptor for dickkopf protein. Cooperates with Dickkopf
 CC to block Wnt/beta-catenin signaling. Forms a ternary complex with
 CC Dkk1 and LRP6 and induces rapid endocytosis and removal of the Wnt
 CC receptor LRP6 from the plasma membrane.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -!- SIMILARITY: Contains 1 CUB domain.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 WSC domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 DR EMBL; AJ457192; CAD29805.1; -.
 DR MGD; MGI:1920266; Kremen2.
 DR InterPro: IPR000859; CUB.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR002883; WSC.
 DR Pfam: PF00431; CUB; 1.
 DR Pfam: PF00051; Kringle; 1.
 DR Pfam: PF01822; WSC; 1.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00042; CUB; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00321; WSC; 1.
 DR PROSITE: PS01180; CUB; 1.
 DR PROSITE: PS00021; KRINGLE 1; 1.
 DR PROSITE: PS00070; KRINGLE 2; 1.
 KW Wnt signaling pathway; Glycoprotein; Kringle; Signal; Transmembrane.
 FT SIGNAL 1 24
 FT CHAIN 25 461
 FT DOMAIN 25 363
 FT TRANSMEM 364 386
 FT DOMAIN 387 461
 FT DOMAIN 34 118
 FT DOMAIN 120 214
 FT DOMAIN 218 325
 FT CARBOHYD 48 48
 FT CARBOHYD 221 221
 FT CARBOHYD 243 243
 FT CARBOHYD 350 350
 SQ SEQUENCE 461 AA; 49170 MW; 6D58C4A2858E09DA CRC64;
 Query Match 17.5%; Score 146.5; DB 1; Length 461;
 Best Local Similarity 42.0%; Pred. No. 5.8e-07;
 Matches 29; Conservative 8; Mismatches 27; Indels 5; Gaps 2;
 QY 50 CYEGNGHFYRGK---ASTDTMGRPCLPNSATVLOQTYHAHRSDALQLGLGKHNYCRNP 106
 DB 35 CFQVNGADYRGHONYTGPGRGAPCLFWDQTC--QHSYSSASDPQGRWGLGAHFCRNP 92
 QY 107 NRRRPWCYV 115
 DB 93 GDVQPCYV 101

DR	ENBL; D90330; BAA14348.1; JOINED.
DR	ENBL; D90331; BAA14348.1; JOINED.
DR	ENBL; D90332; BAA14348.1; JOINED.
DR	ENBL; D90333; BAA14348.1; JOINED.
DR	EMBL; M29145; AAA52650.1; -.
DR	EMBL; M60718; AAA52648.1; -.
DR	EMBL; X16323; CAA34387.1; -.
DR	EMBL; W73239; AAJ64239.1; -.
DR	EMBL; W73240; AAJ64297.1; -.
DR	EMBL; AY246560; AAC61091.1; -.
DR	EMBL; AC004960; AAC71655.1; -.
DR	EMBL; W75983; AAGS3460.1; -.
DR	EMBL; W75972; AAGS3460.1; JOINED.
DR	EMBL; W75973; AAGS3460.1; JOINED.
DR	EMBL; W75974; AAGS3460.1; JOINED.
DR	EMBL; W75975; AAGS3460.1; JOINED.
DR	EMBL; W75976; AAGS3460.1; JOINED.
DR	EMBL; W75977; AAGS3460.1; JOINED.
DR	EMBL; W75978; AAGS3460.1; JOINED.
DR	EMBL; W75979; AAGS3460.1; JOINED.
DR	EMBL; W75980; AAGS3460.1; JOINED.
DR	EMBL; W75981; AAGS3460.1; JOINED.
DR	EMBL; W75982; AAGS3460.1; JOINED.
DR	PIR; JH0579; JH0579.
DR	PDB; 2HGCF; 24-JUN-98.
DR	PDB; 1BHT; 18-NOV-98.
DR	PDB; INKI; 29-DEC-99.
DR	MEROPE; S01.976; -.
DR	GlycoSuiteDB; P14210; -.
DR	Genew; HGNC:4893; HGF.
DR	NIM; 142409; -.
DR	GO; GO:0008083; P:Growth factor activity; NAS.
DR	GO; GO:0007067; P:mitosis; NAS.
DR	InterPro; IPR009003; Cys_Ser_trypsin.
DR	InterPro; IPR000001; Kringie.
DR	InterPro; IPR003014; PAN.
DR	InterPro; IPR003609; Pan_app.
DR	InterPro; IPR001254; Peptidase_S1.
DR	InterPro; IPR001314; Peptidase_SIA.
DR	Pfam; PF00051; kringie; 4.
DR	Pfam; PF00024; PAN; 1.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	PRINTS; PR00018; KRINGLE; 4.
DR	ProDom; PD000395; Kringie; 4.
DR	SMART; SM00130; KR; 4.
DR	SMART; SM00473; PAN_AP; 1.
DR	SMART; SM00020; Tryp_SPC; 1.
DR	PROSITE; PS00021; KRINGLE_1; 4.
DR	PROSITE; PS00070; KRINGLE_2; 4.
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.
KW	Growth factor; Kringie; Glycoprotein; Serine protease homolog; Repeat;
KW	Signal; 3D-structure; Polymorphism; Pyrolidone carboxylic acid.
FT	SIGNAL 1 31
FT	CHAIN 32 494
FT	CHAIN 495 728
FT	MOD_RES 32 32
FT	DOMAIN 32 127
Q	EV Watch 17.3%; Score 144.5; DB 1; Length 728;
I	Local Similarity 29.6%; Pred.No.1.4e-06;
I	ches 45; Conservative 13; Mismatches 69; Indels 25; Gaps 10;
QY	2 NELHQVPSCDC--LNGGTCSVKSYFNIHC-----NCPKFGGO--HCEIDKSKTCYE 52
Dd	335 HEHDMTPENFKDLRENYC-RNPDGSESFWCTFDPNIRVGYSQIPNCDMSHGQDCVR 393
QY	53 GNGHFYRGKASDTDMGRCLPWSATVLQTYIAH---RSDALQLGLGKHNYCRNP-DNR 108
Dd	394 GNKKYMGNLQTSRGLTCSMWDXN---MEDLRHRIFWEPDASKL---NENYCRNPDDDA 447
QY	109 RPPWCIVQVGLKLPLV--QECMVHDCADGKKPS 138

InterPro	IPR000001	Kringle.
InterPro	IPR003014	PAN.
InterPro	IPR003609	Pan_app.
InterPro	IPR001254	Peptidase_S1.
InterPro	IPR001314	Peptidase_S1A.
Pfam	PF00051	kringle; 4.
Pfam	PF00024	PAN; 1.
Pfam	PF00089	trypsin; 1.
PRINTS	PR00722	CHYMOTRYPSIN.
PRINTS	PR00018	KRINGLE.
ProDom	PD000395	Kringle; 4.
SMART	SM00130	KR; 4.
SMART	SM00473	PAN AP; 1.
SMART	SM00020	Tryp_SPC; 1.
PROSITE	PS00021	KRINGLE_1; 4.
PROSITE	PS0070	KRINGLE_2; 4.
PROSITE	PS0240	TRYPSIN_DOM; 1.
Growth factor; Kringle; Glycoprotein; Serine protease homology; Repeat; Signal; Alternative splicing; Pyridone carboxylic acid.		
CHAIN	33	495
CHAIN	496	728
MOD_RES	33	33
DOMAIN	33	128
DOMAIN	129	207
DOMAIN	212	289
DOMAIN	306	384
DOMAIN	392	470
DOMAIN	496	728
DISULFID	71	97
DISULFID	75	85
CARBOHYD	295	295
CARBOHYD	403	403
CARBOHYD	569	569
CARBOHYD	656	656
VARSPLIC	163	167
CONFLICT	344	344
CONFLICT	479	479
CONFLICT	564	564
SEQUENCE	728 AA	82944 MW; A0381FC497534328 CRC64;
very Match	17.0%;	Score 142; DB 1; Length 728;
Local Similarity	28.2%;	Pred.No. 2.5e-06;
atches 44;	Conservative 12;	Mismatches 62; Indels 38; Gaps 11;
Q	5 HOV-PSNCCC--LNGGTCVSNKYFSLHWC-----NCPKFGQCHCEIDKSK 48	
D	338 HDITPENFKDLRENYC-RNPGASPCWFTDPNIRVGYCSQIPK-----CDVSSGQ 390	
Q	49 TCYEGNGHFYRGKASTDTMRGCLPWNASATVLQOTYHAH---RSDALQLGLGKINYCRNP 105	
D	391 DCYRNGKNGVGNLSKTRSGLTCSMWDKN---MEDLHRHIFWEPASKL---NKNYCRNP 444	
Q	106 -DNERPNCYQVGLKPLV--QECVHDCADGKKPS 138	
D	445 DDDAGHPWCYTG---NPLIPWDYCPISRCBGDTTPT 477	
R	JLT 46	
I	HGF_RAT	STANDARD; PRT; 728 AA.
A	P17945;	
D	01-NOV-1990 (Rel. 16, Created)	
D	01-NOV-1990 (Rel. 16, Last sequence update)	
D	28-FEB-2003 (Rel. 41, Last annotation update)	
D	Hepatocyte growth factor precursor (Scatter factor) (SF)	
C	(Hepatopoietin A).	
C	Rattus norvegicus (Rat).	
C	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116;

[1]

SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RP STRAIN=Wistar; TISSUE=Liver;

RC MEDLINE=90222197; PubMed=2139229;

RX Toshiro K., Hagiya M., Nishizawa T., Seki T., Shimonishi M., Shimizu S., Nakamura T.;

RA "deduced primary structure of rat hepatocyte growth factor and expression of the mRNA in rat tissues.;"

RT Proc. Natl. Acad. Sci. U.S.A. 87:3200-3204(1990).

RL [2]

SEQUENCE FROM N.A.

RP STRAIN=Wistar; TISSUE=Liver;

RC MEDLINE=91031482; PubMed=2146117;

RX Okajima A., Miyazawa K., Kitamura N.;

RA "Primary structure of rat hepatocyte growth factor and induction of its mRNA during liver regeneration following hepatic injury.;"

RT Eur. J. Biochem. 193:375-381(1990).

CC -!- FUNCTION: HGF is a potent mitogen for mature parenchymal hepatocyte cells, seems to be an hepatotrophic factor, and acts as growth factor for a broad spectrum of tissues and cell types. It has no detectable protease activity.

CC -!- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a disulfide bond.

CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.

CC -!- SIMILARITY: Contains 4 kringle domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; D90102; BAAL4133.1; -.

DR EMBL; X54400; CAA38266.1; -.

DR PIR; A35644; A35644.

DR HSP; P14210; 1BHT.

DR MEROPS; S01.978; -.

DR InterPro; IPR003003; Cys_ser_trypsin.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003014; PAN.

DR InterPro; IPR003609; Pan app.

DR InterPro; IPR001254; Peptidase S1.

DR InterPro; IPR001314; Peptidase_S1A.

DR Pfam; PF00051; kringle; 4.

DR Pfam; PF00024; PAN; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYNOTRYPsin.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 4.

DR SMART; SM00130; KR; 4.

DR SMART; SM00473; PAN_AP; 1.

DR SMART; SM00020; TRYp_SPC; 1.

DR PROSITE; PS00021; KRINGLE_1; 4.

DR PROSITE; PS00070; KRINGLE_2; 4.

DR PROSITE; PS00240; TRYPSIN_DOM; 1.

CC Growth factor; Kringle; Glycoprotein; Serine protease homolog; Repeat; Signal; Pyrrolidone carboxylic acid.

FT SIGNAL 1 22 BY SIMILARITY

FT CHAIN 33 485 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.

FT CHAIN 496 728 HEPATOCYTE GROWTH FACTOR BETA CHAIN.

FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID (BY SIMILARITY).

FT FT

FT DOMAIN 33 128 PAP.

FT DOMAIN 129 207 KRINGLE 1.

FT DOMAIN 212 289 KRINGLE 2.

FT DOMAIN 306 384 KRINGLE 3.

FT DOMAIN 392 470 KRINGLE 4.

FT DOMAIN 496 728 SERINE PROTEASE-LIKE.

FT DISULFID 71 97 BY SIMILARITY.

DISULFID 75 85 BY SIMILARITY.
 PT DISULFID 488 607 INTERCHAIN (BY SIMILARITY).
 PT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 PT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 656 656 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 728 AA; 82905 MW; 3E0BF1F96ADCEDFF CRC64;
 Query Match 17.0%; Score 142; DB 1; Length 728;
 Best Local Similarity 26.8%; Pred. No. 2,5e-06;
 Matches 45; Conservative 11; Mismatches 62; Indels 38; Gaps 11;
 QY 5 HQV-FSNDC--LNGGTCVSNKFNHWC-----NCPKFGQGHCEIDKSK 48
 DB 338 HDITENFKDLRENYC-RNPDGAEPCWCTDPNIRVGYSQIPK-----CDVSSGQ 390
 QY 49 TCVEGNHGYRCKASTDTMGRCPLPNSATVLQOYTHAH---RSDALGLGKNGYCRNP 105
 DB 391 DCYRGNKNGYMLKSTRSGLTCSWDKN---MEDLHRHIFWEPDASKL---TKNYCRNP 444
 QY 106 -DNRERPWCYVQVGLKPLV--QECMVHDCADGKPS 138
 DB 445 DDDAGHPWCYTG---NPLVPWDCYPSRCGDTTPT 477
 RE PETMA
 PL PLMN PETMA STANDARD; PRT; 325 AA.
 AC P33574;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasminogen (EC 3.4.21.7) (Fragments).
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE
 RA Affolter M., Schaller J., Rickli E.E.;
 RT Isolation, characterization and partial amino acid sequence of
 RL lamprey plasminogen.";
 RL Protein Seq. Data Anal. 5:207-211(1993).
 CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
 CC a proteolytic factor in a variety of other processes including
 CC embryonic development, tissue remodeling, tumor invasion,
 CC and inflammation.
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
 CC -!- SIMILARITY: Contains at least 2 kringle domains.
 DR PIR; S33879; S33879.
 DR HSP; P00747; SHPG.
 DR MEROPS; S01.233; .
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase_S1.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00021; KRINGLE1; 2.
 DR PROSITE; PS00070; KRINGLE2; 2.
 DR PROSITE; PS0240; TRYPSIN_DOM; PARTIAL.
 DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
 DR PROSITE; PS00135; TRYPSIN_SSR; PARTIAL.
 DR Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Zymogen.
 FT NON_CONS 15 16
 FT NON_CONS 34 45
 FT NON_CONS 44 45
 FT NON_CONS 76 77

PT NON_CONS 111 112
 PT NON_CONS 138 139
 PT NON_CONS 158 159
 FT NON_CONS 178 179
 FT NON_CONS 216 217
 FT NON_CONS 236 237
 FT NON_CONS 267 268
 FT NON_CONS 282 283
 FT NON_CONS 295 296
 FT NON_CONS 307 308
 FT NON_CONS 315 316
 FT NON_TER 325 325
 SQ SEQUENCE 325 AA; 35194 MW; 1B5F0B539AC6ED3C CRC64;
 Query Match 16.2%; Score 136; DB 1; Length 325;
 Best Local Similarity 30.7%; Pred. No. 4,3e-06;
 Matches 31; Conservative 7; Mismatches 41; Indels 22; Gaps 3;
 QY 34 PKFGQGHCEIDKSKTCYEGNGHFGYRKGASTDTMGRCPLPNSATVLQOYTHAHSALQ 93
 DB 71 POSFAG-----LTACVKGTEGVRGTAALTAVSGKACQAWASQT-----PDVYS 115
 QY 94 LGLGKNGYCRNPDRRPPWCYVQVGLKPLVQECMVHDCADG 134
 DB 116 CQGLVSNYCRNPDEKLPWCYT-----TEYCNVPSCTGG 149
 RESULT 48
 THRE MOUSE
 ID THRE MOUSE STANDARD; PRT; 618 AA.
 AC P19221;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Prothrombin precursor (EC 3.4.21.5).
 DE F2 OR CP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Liver;
 RX MEDLINE=9102551; PubMed=222810;
 RX Friezner Degen S.J., Schaffer L.A., Jamison C.S., Grant S.G.,
 RA Fitzgibbon J.J., Pai J.-A., Chapman V.M., Elliott R.W.;
 RA "Characterization of the cDNA coding for mouse prothrombin and
 RT localization of the gene on mouse chromosome 2.";
 RL DNA Cell Biol. 9:487-498(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Liver;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Suetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Capleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schurch A., Schein J.E., Jones S.D.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[3] SEQUENCE OF 384-618 FROM N.A.
 TISSUE=Liver;
 MEDLINE=92212913; PubMed=1557383;
 Banfield D.K., Macgillivray R.R.;
 "Partial characterization of vertebrate prothrombin cDNAs:
 amplification and sequence analysis of the B chain of thrombin from
 nine different species.";
 Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
 C -!- FUNCTION: Thrombin, which cleaves bonds after Arg and Lys,
 converts fibrinogen to fibrin and activates factors V, VII, VIII,
 XI, and, in complex with thrombomodulin, protein C.
 C -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
 fibrinogen to fibrin and releases fibrinopeptide A and B.
 C -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
 RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
 ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
 ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
 CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION
 OF PROTHROMBIN TO THROMBIN.
 C -!- MISCELLANEOUS: Prothrombin is activated on the surface of a
 phospholipid membrane that binds the amino end of prothrombin and
 factors Va and Xa in Ca-dependent interactions; factor Xa removes
 the activation peptide and cleaves the remaining part into light
 and heavy chains. The activation process starts slowly because
 factor V itself has to be activated by the initial, small amounts
 of thrombin.
 C -!- MISCELLANEOUS: Thrombin can itself cleave the amino terminal
 fragment (fragment 1) of the prothrombin, prior to its activation
 by factor Xa.
 C -!- SIMILARITY: Belongs to peptidase family S1.
 C -!- SIMILARITY: Contains 2 kringle domains.
 C This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 D EMBL; X52308; CAA36548.1; -
 D EMBL; BC013662; AAH13662.1; -
 D EMBL; M81394; AAA40435.1; -
 D F1; A35827; A35827.
 D HSP; P00734; I87X.
 D MEROPS; S01.217; -
 D InterPro; IPR009003; Cys_Ser_trypsin.
 D InterPro; IPR002383; GLA_blood.
 D InterPro; IPR000001; Kringle.
 D InterPro; IPR001254; Peptidase_S1.
 D InterPro; IPR001314; Peptidase_S1A.
 D InterPro; IPR003966; Peptidase_S1A_pr.
 D InterPro; IPR000294; Vitk_dep_GLA.
 D Pfam; PF00594; gla; 1.
 D Pfam; PF00051; kringle; 2.
 D Pfam; PF00089; trypsin; 1.
 D PRINTS; PR00722; CHYMOTRYPSIN.
 D PRINTS; PR00001; GLABLOOD.
 D PRINTS; PR00018; KRINGLE.
 D PRINTS; PR01505; PROTHROMBIN.
 D ProDom; PD000395; Kringle; 2.
 D SMART; SM00069; GLA; 1.
 D SMART; SM00130; KR; 2.
 D SMART; SM00020; TRYPSIN; 1.
 D PROSITE; PS00011; GLI_CARBOXYLATION; 1.
 D PROSITE; PS00021; KRINGLE_1; 2.
 D PROSITE; PS00070; KRINGLE_2; 2.
 D PROSITE; PS00240; TRYPSIN_DOM; 1.
 D PROSITE; PS00134; TRYPSIN_HIS; 1.
 D PROSITE; PS00135; TRYPSIN_SER; 1.
 D Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;

KW Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
 KW Hydrolyase; Serine protease; Kringle; Signal.
 FT SIGNAL 1 24
 FT PROPEP 25 43
 FT CHAIN 44 618
 FT PEPTIDE 44 200 PROTHROMBIN
 FT PEPTIDE 201 324 ACTIVATION PEPTIDE (FRAGMENT 1).
 FT CHAIN 325 360 ACTIVATION PEPTIDE (FRAGMENT 2).
 FT CHAIN 361 618 THROMBIN LIGHT CHAIN (A).
 FT CHAIN 619 187 THROMBIN HEAVY CHAIN (B).
 FT DOMAIN 109 187 KRINGLE 1.
 FT DOMAIN 215 292 KRINGLE 2.
 FT DOMAIN 361 618 SERINE PROTEASE.
 FT SITE 200 201 CLEAVAGE (BY THROMBIN).
 FT SITE 324 325 CLEAVAGE (BY FACTOR XA).
 FT SITE 360 361 CLEAVAGE (BY FACTOR XA).
 FT ACT_SITE 403 403 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 459 459 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 565 565 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT MOD_RES 50 50 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 51 51 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 58 58 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 63 63 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 64 64 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 73 73 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 76 76 GAMMA-CARBOXYGLUTAMIC ACID.
 FT DISULFID 61 66 BY SIMILARITY.
 FT DISULFID 91 104 BY SIMILARITY.
 FT DISULFID 109 187 BY SIMILARITY.
 FT DISULFID 130 170 BY SIMILARITY.
 FT DISULFID 158 182 BY SIMILARITY.
 FT DISULFID 215 293 BY SIMILARITY.
 FT DISULFID 236 276 BY SIMILARITY.
 FT DISULFID 264 288 BY SIMILARITY.
 FT DISULFID 333 479 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 388 404 BY SIMILARITY.
 FT DISULFID 533 547 BY SIMILARITY.
 FT DISULFID 561 591 BY SIMILARITY.
 FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .).
 SQ SEQUENCE 618 AA; 70268 MW; B89F719AAPD601E0 CRC64;
 Query Match 15.0%; Score 125.5; DB 1; Length 618;
 Best Local Similarity 31.8%; Pred. No. 8.6e-05;
 Matches 27; Conservative 13; Mismatches 40; Indels 5; Gaps 3;
 Qy 50 CYEGNGHFYRGKASTDTMGPCPLPWSATVLTQTYHAHRSALQLGLGKHNYCRNPD-NR 108
 Db 215 CLTERGLYQGNLAVITLGSPLPWSLPAKTLISKYQDFPEVKL---VENFCRNFWDWE 271
 Qy 109 RRPWCYVQGLKPLVQECMVHDCAD 133
 Db 272 EGAWCYV-AGQPGDFEYCNLYCEE 295
 RESULT 49
 ID PLMN_SHEEP STANDARD; PRT; 343 AA.
 AC P81286;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasminogen (EC 3.4.21.7) (Fragment).
 GN PLG.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OC NCBI_TaxID=9940;

```

[1]
RN RP MEDLINE=93149995; PubMed=1492092;
SEQUENCE.
RR Schaller J., Straub C., Kamper U., Rickli E.E.;
RT "Complete amino acid sequence of ovine miniplasminogen.";
RL Protein Seq. Data Anal. 5:21-25(1992).
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
CC a proteolytic factor in a variety of other processes including
CC embryonic development, tissue remodeling, tumor invasion, and
CC inflammation; in ovulation it weakens the walls of the Graafian
CC follicle. It activates the urokinase-type plasminogen activator,
CC collagenases and several complement zymogens, such as C1 and C5.
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von
CC Willebrand factor.
CC CC
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC CC
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen
CC activators, both plasminogen and its activator being bound to
CC fibrin. Cannot be activated with streptokinase.
CC CC
CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
CC InterPro; IPR0011254; Peptidase S1.
CC InterPro; IPR001314; Peptidase S1A.
CC InterPro; IPR003966; Peptidase_S1A_pr.
CC Pfam; PF00051; Kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PRO00722; CHYMOTRYPSIN.
CC PRINTS; PRO00018; KRINGLE.
CC PRINTS; PRO1505; PROTHROMBIN.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp Spc; 1.
CC PROSITE; PS00021; KRINGLE-1; 1.
CC PROSITE; PS00070; KRINGLE-2; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
CC Tissue remodeling; blood coagulation; Kringle; Zymogen; Repeat.
CC FT NON_TER 1 1
CC FT DOMAIN <1 140 HEAVY CHAIN A.
CC FT DOMAIN 141 >343 LIGHT CHAIN A.
CC FT DOMAIN <1 17 KRINGLE 4.
CC FT DOMAIN 41 120 KRINGLE 5.
CC FT DOMAIN 114 341 SERINE PROTEASE.
CC FT ACT_SITE 181 181 CHARGE RELAY SYSTEM.
CC FT ACT_SITE 224 224 CHARGE RELAY SYSTEM.
CC FT ACT_SITE 319 319 CHARGE RELAY SYSTEM.
CC FT NON_TER 343 343
CC SEQUENCE 343 AA; 37662 MW; 8DF6BEA92D596EE0 CRC64;
RY Watch 14.9%; Score 124.5; DB 1; Length 343;
Local Similarity 35.3%; Pred.No.5.9e-05;
ches 36; Conservative 8; Mismatches 43; Indels 15; Gaps 6;
QY 50 CYEGNGHYRGKASTDTWGRCPLPWN SATVLTQTYTAH-----RSDALQLGKGNHCNP 105
DB 15 CMLGIGKGRGKATVAGVPCQEWAA-----QEPHRHGIFTETNPAGLSEK-NVCNP 68
QY 106 D-NRRRPWCYTVGLKPLVQECMVHDC-----DGKKSPSPPE 143
DB 69 GDGVNPGWCYT-TNPKLFDYCDIPQCESSFDCKPKEPDK 109
JT 50

```

R R X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
R MEDLINE=97357286; PubMed=9214615;
R van de Loch A., Bode W., Huber R., le Bonniec B.F., Stone S.R.,
R Esmen C.T., Stubbs M.T.;
R "The thrombin E192Q-BPTI complex reveals gross structural
R rearrangements: implications for the interaction with antithrombin
R and thrombomodulin.";
R EMBO J. 16:2977-2984(1997).
R [11]
R X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 328-601.
R MEDLINE=99162521; PubMed=10051558;
R Quinto E.R., Caccia S., Rose T., Fueterer K., Waksman G., di Cera E.;
R "Unexpected crucial role of residue 225 in serine proteases.";
R Proc. Natl. Acad. Sci. U.S.A. 96:1852-1857(1999).
R [12]
R VARIANT DYSPROTHROMBINEMIA CYS-314.
R MEDLINE=87033739; PubMed=3771562;
R Rabiet M.-J., Furlie B.C., Furlie B.;
R "Molecular defect of prothrombin Barcelona. Substitution of cysteine
R for arginine at residue 273.";
R J. Biol. Chem. 261:15048-15048(1986).
R [13]
R VARIANT DYSPROTHROMBINEMIA ALA-509.
R MEDLINE=95313001; PubMed=7792730;
R Degen S.J.P., McDowell S.A., Sparks L.M., Scharrer I.;
R "Prothrombin: Frankfurt: a dysfunctional prothrombin characterized by
R substitution of Glu-466 by Ala.";
R Thromb. Haemost. 73:203-209(1995).
R [14]
R VARIANTS DYSPROTHROMBINEMIA THR-380 AND HIS-431.
R MEDLINE=93043342; PubMed=1421398;
R Morishita E., Saito M., Kumabashiri I., Asakura H., Matsuda T.,
R Yamaguchi K.;
R "Prothrombin.Himi: a compound heterozygote for two dysfunctional
R prothrombin molecules (Met-337-->Thr and Arg-388-->His).";
R Blood 80:2275-2280(1992).
R [15]
R VARIANT DYSPROTHROMBINEMIA HIS-314.
R MEDLINE=95169898; PubMed=7865694;
R James H.L., Kim D.J., Zheng D.-Q., Girolami A.;
R "Prothrombin: Padua I: incomplete activation due to an amino acid
R substitution at a factor Xa cleavage site.";
R Blood Coagul. Fibrinolysis 5:841-844(1994).
R [16]
R VARIANT DYSPROTHROMBINEMIA CYS-425.
R MEDLINE=89207504; PubMed=3242619;
R Henriksen R.A., Mann K.G.;
R "Identification of the primary structural defect in the dysfibrinogen
R thrombin Quick I: substitution of cysteine for arginine-382.";
R Biochemistry 27:9160-9165(1988).
R [17]
R VARIANT DYSPROTHROMBINEMIA VAL-601.
R MEDLINE=89247398; PubMed=2719946;
R Henriksen R.A., Mann K.G.;
R "Substitution of valine for glycine-558 in the congenital dysfibrinogen
R thrombin Quick II alters primary substrate specificity.";
R Biochemistry 28:2078-2082(1989).
R [18]
R VARIANT DYSPROTHROMBINEMIA ALA-509.
R MEDLINE=92378975; PubMed=1354985;
R Miyata T., Aruga R., Umayama H., Bezeaud A., Guillin M.-C.,
R Iwanaga S.;
R "Prothrombin Salakta: substitution of glutamic acid-466 by alanine
R reduces the fibrinogen clotting activity and the esterase activity.";
R Biochemistry 31:7457-7462(1992).
R [19]
R VARIANT DYSPROTHROMBINEMIA TRP-461.
R MEDLINE=87185407; PubMed=3567158;
R Miyata T., Morita T., Inomoto T., Kawauchi S., Shirakami A.,
R Iwanaga S.;
R "Prothrombin Tokushima, a replacement of arginine-418 by tryptophan
R that impairs the fibrinogen clotting activity of derived thrombin
R Tokushima.";

RL Biochemistry 26:1117-1122(1987).
RN [20]
RX VARIANT DYSPROTHROMBINEMIA TRP-461.
RX MEDLINE=87101511; PubMed=3801671;
RA Inomoto T., Shirakami A., Kawauchi S., Shigekiyo T., Saito S.,
RA Miyoshi K., Morita T., Iwanaga S.;
RT "Prothrombin Tokushima: characterization of dysfunctional thrombin
RT derived from a variant of human prothrombin.";
RL Blood 69:565-569(1987).
RN [21]
RN VARIANT DYSPROTHROMBINEMIA TRP-461.
RX MEDLINE=92256895; PubMed=1349838;
RA Iwahana H., Yoshimoto K., Shigekiyo T., Shirakami A., Saito S.,
RA Itakura M.;
RT "Detection of a single base substitution of the gene for prothrombin
RT Tokushima. The application of PCR-SSCP for the genetic and molecular
RT analysis of dysprothrombinemia.";
RL Int. J. Hematol. 55:93-100(1992).
RN [22]
RN VARIANT DYSPROTHROMBINEMIA LYS-200.
RX MEDLINE=83204687; PubMed=6405779;
RA Board P.G., Shaw D.C.;
RT "Determination of the amino acid substitution in human prothrombin
RT type 3 (157 Glu leads to Lys) and the localization of a third
RT thrombin cleavage site.";
RL Bx. J. Haematol. 54:245-254(1983).
RN [23]
RN VARIANTS MET-165 AND THR-386.
RX MEDLINE=99318093; PubMed=10391209;
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
RT of human genes.";
RL Nat. Genet. 22:231-238(1999).
RN [24]
RN ERRATUM.
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RL Nat. Genet. 23:373-373(1999).
CC -1- FUNCTION: Thrombin, which cleaves bonds after Arg and Lys,
CC converts fibrinogen to fibrin and activates factors V, VII, VIII,
CC XIII, and, in complex with thrombomodulin, protein C.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-Gly; activates
CC fibrinogen to fibrin and releases fibrinopeptide A and B.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC ENZYME, THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS.
CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE, THE MODIFIED RESIDUES
CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION

Query Match 14.8%; Score 124; DB 1; Length 622;
Best Local Similarity 38.8%; Pred. No. 0.00012;
Matches 26; Conservative 8; Mismatches 29; Indels 4; Gaps 2;
Qy 50 CYSNGHGFYRGKASTDTMGSPCLPWNASATVLTQTYHAHSDALQGLGKHNYCEND-NR 108
Db 213 CVPRGQYQGRILAVTTHTGLPLCLAWASQAQKALSKHDFNSAVQL---VENFCRNPDGDE 269
Qy 109 RRPWCYV 115
Db 270 EGVCYV 276

Search completed: May 25, 2004, 14:53:56
Job time : 6.51446 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

rotein - protein search, using sw model

on: May 25, 2004, 14:44:05 ; Search time 23.6871 Seconds
(without alignments)
1904.795 Million cell updates/sec

e: US-09-880-503-8

ect score: 837

ence: 1 SNELHQVPSNCDLNGTGV.....QECMVHDCADKKPSPPEE 143

ing table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 1017041 seqs, 315518202 residues

il number of hits satisfying chosen parameters: 1017041

imum DB seq length: 0

imum DB seq length: 2000000000

rocessing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

base :

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

R	lt	Score	Query Match	Length	DB ID	Description
1	778	93.0	154	4	Q96SE8	Q96SE8 homo sapien
2	627	74.9	433	6	Q8MI10	Q8MI10 oryctolagus
3	627	74.9	433	6	Q8MHY7	Q8MHY7 oryctolagus
4	576	68.8	157	6	Q9TVA8	Q9TVA8 bos taurus
5	569	68.0	231	11	Q8C6L2	Q8C6L2 mus musculus
6	337.5	40.3	214	6	Q9X170	Q9X170 oryctolagus
7	328.5	39.2	291	4	Q7Z7N2	Q7Z7N2 homo sapien
8	328.5	39.2	562	4	Q86YK8	Q86YK8 homo sapien
9	322.5	38.5	516	4	Q9BU99	Q9BU99 homo sapien
10	311.5	37.2	562	6	Q8SQ23	Q8SQ23 sus scrofa
11	301.5	36.0	564	6	Q8MK51	Q8MK51 oryctolagus
12	265	31.7	517	11	Q8K0D2	Q8K0D2 mus musculus
13	263.5	31.5	560	4	Q14520	Q14520 homo sapien
14	258	30.8	616	6	Q97507	Q97507 sus scrofa
15	255.5	30.5	540	13	Q800Y7	Q800Y7 meleagris g
16	250.5	29.9	653	11	Q8VCS4	Q8VCS4 mus musculus

ALIGNMENTS

RESULT 1	Q96SE8	PRELIMINARY;	PRT;	154 AA.
Q96SE8	AC	Q96SE8		
DT	01-DEC-2001	(TrEMBLrel. 19, Created)		
DT	01-DEC-2001	(TrEMBLrel. 19, Last sequence update)		
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)		
DE	Urokinase-type plasminogen activator amino-terminal fragment.			
GN	ATF.			

17	250	29.9	597	11	Q35727	Q35727 mus musculus
18	250	29.9	609	11	Q80VC5	Q80VC5 mus musculus
19	233	27.8	395	4	Q8ZM1	Q8ZM1 homo sapien
20	233	27.8	615	4	Q8IZZ5	Q8IZZ5 homo sapien
21	218.5	26.1	90	4	Q8NG20	Q8NG20 homo sapien
22	197.5	23.6	202	13	Q90675	Q90675 gallus gall
23	174	20.8	421	13	Q8AXX3	Q8AXX3 xenopus lae
24	174	20.8	947	13	Q8AXY6	Q8AXY6 gallus gall
25	166	19.8	263	13	Q7SXB3	Q7SXB3 brachydanio
26	166	19.8	716	13	Q91691	Q91691 xenopus lae
27	165.5	19.8	359	6	Q8WML1	Q8WML1 canis famil
28	165.5	19.8	812	11	Q9ROW3	Q9ROW3 rattus norv
29	158.5	18.9	393	4	Q9BRB6	Q9BRB6 homo sapien
30	158.5	18.9	937	11	Q8BNP9	Q8BNP9 mus musculus
31	158.5	18.9	937	11	Q8BG10	Q8BG10 mus musculus
32	158	18.9	454	6	Q46506	Q46506 papio hamad
33	157	18.8	113	4	Q9UIR5	Q9UIR5 homo sapien
34	156.5	18.7	759	11	Q7TP84	Q7TP84 rattus norv
35	156.5	18.7	810	4	Q51446	Q51446 homo sapien
36	154	18.4	801	11	Q8KQ08	Q8KQ08 mus musculus
37	154	18.4	806	6	Q18783	Q18783 macropus eu
38	154	18.4	944	11	Q8C3W2	Q8C3W2 mus musculus
39	154	18.4	944	11	Q8BGP6	Q8BGP6 mus musculus
40	153	18.3	113	4	Q9UIR7	Q9UIR7 homo sapien
41	152.5	18.2	717	13	Q700C6	Q700C6 xenopus lae
42	152	18.2	567	4	Q3208	Q3208 homo sapien
43	152	18.2	648	4	Q9H1V4	Q9H1V4 homo sapien
44	150.5	18.0	930	13	Q8AV69	Q8AV69 xenopus lae
45	149.5	17.9	709	13	Q7ZTN9	Q7ZTN9 xenopus lae
46	148	17.7	105	4	Q9UIR8	Q9UIR8 homo sapien
47	148	17.7	704	13	Q90865	Q90865 gallus gall
48	147.5	17.6	716	11	Q91XG8	Q91XG8 mus musculus
49	147	17.6	113	4	Q9UIR6	Q9UIR6 homo sapien
50	147	17.6	2869	6	Q28398	Q28398 erinaceus e
51	146.5	17.5	709	13	Q902N6	Q902N6 brachydanio
52	146.5	17.5	902	5	Q17576	Q17576 caenorhabdi
53	146.5	17.5	928	5	Q9BLX1	Q9BLX1 caenorhabdi
54	145.5	17.4	313	13	Q9PU78	Q9PU78 crocodylus
55	145.5	17.4	716	11	Q70521	Q70521 rattus norv
56	145	17.3	726	13	Q90978	Q90978 gallus gall
57	142.5	17.0	213	11	Q81123	Q81123 mus musculus
58	142.5	17.0	264	11	Q81122	Q81122 mus musculus
59	142.5	17.0	264	11	Q7TMJ8	Q7TMJ8 mus musculus
60	142	17.0	215	13	Q42341	Q42341 gallus gall
61	142	17.0	728	6	Q9BH09	Q9BH09 felis silve
62	142	17.0	728	11	Q8C9G5	Q8C9G5 mus musculus
63	142	17.0	730	6	Q867B7	Q867B7 canis famil
64	142	17.0	948	5	Q911Y6	Q911Y6 ephydatia f
65	141	16.8	710	13	Q91402	Q91402 xenopus. he
66	140	16.7	132	4	Q16609	Q16609 homo sapien
67	138	16.5	1145	5	Q9BK18	Q9BK18 aplysia cal
68	137	16.4	334	6	Q46507	Q46507 papio hamad
69	135.5	16.2	553	13	Q7Z292	Q7Z292 brachydanio
70	135	16.1	429	13	Q8AVD0	Q8AVD0 brachydanio
71	134	16.0	211	11	Q55027	Q55027 mus musculus
72	133.5	15.9	145	6	Q28911	Q28911 macaca fasc
73	131.5	15.7	234	4	Q86YW2	Q86YW2 homo sapien
74	131.5	15.7	263	4	Q00318	Q00318 homo sapien
75	131.5	15.7	263	4	Q8NCJ9	Q8NCJ9 homo sapien

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu J., Bai X., Ruan C.;
RT "Cloning and expression of the amino-terminal fragment of human
RT urokinase-type plasminogen activator."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Overexpression of the amino-terminal fragment of human urokinase-type
RT plasminogen activator in breast cancer cells results in decreased
RT tumor invasion, growth and angiogenesis."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL: AY029537; AAK38734.1;
DR GO: GO:0016301; F.kinase activity; IEA.
DR InterPro: IPR006209; EGF-like.
DR InterPro: IPR000001; Kringle.
DR PRINTS: PR00018; KRINGLE.
DR PRODom: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00051; kringle; 1.
DR PRINTS: PR00018; KRINGLE.
DR PRODom: PD000395; Kringle; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00130; KR; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CCF2FCFF505572 CRC64;

Query Match 93.0%; Score 778; DB 4; Length 154;
Best Local Similarity 99.3%; Pred. No. 1.4e-75;
Matches 133; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNELHQPNCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFRG 60
DB 21 SNELHQPNCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFRG 80

QY 61 KASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQGLK 120
DB 81 KASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQGLK 140

QY 121 PLVQECMVHDCADG 134
DB 141 LLVQECMVHDCADG 154

RESULT 2
ID Q8MHY7 PRELIMINARY; PRT; 433 AA.
AC Q8MHY7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator.
GN PLAU.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22155945; PubMed=12149463;
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA Dichek D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
RT development causes arterial constriction and lumen loss, and
RT accelerates lesion growth."
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC EMBL: AY122285; AAM83187.1;
DR GO: GO:0004263; F.chymotrypsin activity; IEA.
DR GO: GO:0016301; F.kinase activity; IEA.
DR GO: GO:0008233; F.peptidase activity; IEA.
DR GO: GO:0004295; F.trypsin activity; IEA.
DR GO: GO:0006508; P.proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys Ser trypsin.
DR InterPro: IPR006209; EGF-like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Peptidase S1.
DR InterPro: IPR001314; Peptidase S1A.
DR InterPro: IPR008293; Pept S1A_UPA.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00059; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR PRODom: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR PIRSF: PIRSF001144; Urk_plasm_act; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65B64F36415549B0 CRC64;

Query Match 74.9%; Score 627; DB 6; Length 433;
Best Local Similarity 76.2%; Pred. No. 6.8e-59;
Matches 109; Conservative 11; Mismatches 21; Indels 2; Gaps 1;

QY 1 SNELHQV--PSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFRG 58
DB 21 SNELHGVSDASNGCLNGGTCVTVKYFSNIWRCNCPKFGQHCIDTLKTCVHGDSY 80

QY 59 RGASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQVG 118
DB 81 RGKANTDIDMRPCLAWSANVLTKTTHAHRPDALQGLGKHNYCRNPDNRPRPWCYVQVG 140

QY 119 LKPLVQECMVHDCADGCKPSSPP 141
DB 141 LKQLICEKVDHDCSSGCKPALPP 163

RESULT 3
ID Q8MHY7 PRELIMINARY; PRT; 433 AA.
AC Q8MHY7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT complete cds."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

D EMBL; AY029517; AAK40239.1; --
D EMBL; AB087224; BAC02685.1; --
D GO; GO:0004263; F:chymotrypsin activity; IEA.
D GO; GO:0016301; F:kinase activity; IEA.
D GO; GO:0008233; F:peptidase activity; IEA.
D GO; GO:0004295; F:trypsin activity; IEA.
D GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
D InterPro; IPR009003; Cys_Ser_trypsin.
D InterPro; IPR006209; EGF-like.
D InterPro; IPR000001; Kringle.
D InterPro; IPR001254; Peptidase S1.
D InterPro; IPR001314; Peptidase S1A.
D InterPro; IPR008293; Pept_S1A_UPA.
D Pfam; PF00051; kringle; 1.
D Pfam; PF00089; trypsin; 1.
D PRINTS; PR00722; CHYMOTRYPsin.
D PRINTS; PR00018; KRINGLE.
D ProDom; PD000395; Kringle; 1.
D SMART; SM00130; KR; 1.
D SMART; SM00020; Tryp_SPC; 1.
D PROSITE; PS00022; EGF_1; 1.
D PROSITE; PS00021; KRINGLE_1; 1.
D PROSITE; PS00070; KRINGLE_2; 1.
D PROSITE; PS00240; TRYPSIN_DOM; 1.
D PROSITE; PS00134; TRYPSIN_HIS; 1.
D PROSITE; PS00135; TRYPSIN_SER; 1.
D PIRSF; PIRSF001144; Urk_p1asm_act; 1.
D Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
D SEQUENCE 433 AA; 48444 MW; 6D35A371010A6EE CRC64;

Query Match 74.9%; Score 627; DB 6; Length 433;
Best Local Similarity 76.2%; Pred. No. 6.8e-59;
Matches 109; Conservative 11; Mismatches 21; Indels 2; Gaps 1;

Q 1 SNEHQV--PNCDCGLNGCTVSNKYNFNSIHWNCNCPKFGQHCEDKSKTCYEGNGHFY 58
D 21 SHELHGSDASNGCLNGCTVYKYNFNSIWRNCNCPKFGQHCEDTLTKCYHGDSHY 80

Q 59 RKASTDITMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHYCRNPDNRRPWCYVQVG 118
D 81 RKGANTDMDRCLAWNSANVLTITYHAHRPDALQLGLGKHYCRNPDNRRPWCYVQVG 140

Q 119 LKPLVQECWHDCAQKPSPP 141
D 141 LKQLIQECKVHDSSGKKPALUPP 163

R JLT 4
Q /AS
I Q9TVAS PRELIMINARY; PRT; 157 AA.
A Q9TVAS;
D 01-MAY-2000 (TREMBlrel. 13, Created)
D 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
D 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
D Urokinase plasminogen activator (Fragment).
O Bos taurus (Bovine)
O Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
O Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
O Bovidae; Bovinae; Bos.
O NCBI_TaxID=9913;
R [1]
R SEQUENCE FROM N.A.
R TISSUE-Skeletal muscle;
R MEDLINE=21071388; PubMed=11204721;
R Balcerzak D., Quereguesser L., Dixon W.T., Baracos V.E.;
R "Coordinate expression of matrix-degrading proteinases and their
R activators and inhibitors in bovine skeletal muscle.";
R J. Anim. Sci. 79:94-107(2001).
C -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
D EMBL; AF144761; AAD30301.1; --
D HSP; P00749; IURK.
D GO; GO:0016301; F:kinase activity; IEA.
D InterPro; IPR006209; EGF-like.

DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kinase; Kringle.
FT NON_TER 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match 58.8%; Score 576; DB 6; Length 157;
Best Local Similarity 74.8%; Pred. No. 6.7e-54;
Matches 98; Conservative 12; Mismatches 21; Indels 0; Gaps 0;

QY 13 CLNGGTCVSNKYNFNSIHWNCNCPKFGQHCEDKSKTCYEGNGHFYRGRKASTDTMGRPCL 72
DB 1 CLNGGKCVTYKYNFNSIQRCSCEPKFGQHCEDTSKTCYQNGNHSYRGRKANEDLSGRPCL 60

QY 73 PWSATVLOQTYHAHRSDALQLGLGKHYCRNPDNRRPWCYVQVGLKPLVQECWHDCA 132
DB 61 AWDSPTVLLKMYHAHRSDAQLGLGKHYCRNPDNRRPWCYVQVGLKPLVQECWHDCA 120

QY 133 DGKKPSSPPEE 143
DB 121 VGKSPSSPREK 131

RESULT 5
Q8C6L2 PRELIMINARY; PRT; 231 AA.
ID Q8C6L2;
AC Q8C6L2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE-Ovary;
RX MEDLINE=22954683; PubMed=12466851;
RA The RIKEN Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK054349; BAC35743.1; --
DR PIR; PT0534; PT0534.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
FT NON_TER 231 231
SQ SEQUENCE 231 AA; 25510 MW; 25B8980A682737F2 CRC64;

Query Match 58.0%; Score 569; DB 11; Length 231;
Best Local Similarity 70.4%; Pred. No. 5.7e-53;
Matches 95; Conservative 15; Mismatches 25; Indels 0; Gaps 0;


```

QY 9 SNCDLNGTGVSNKYFNSHWCNCPKFGGQHCEIDSKTCYEGNGHGYRKGASTDTMG 68
DB 30 SNGCGQNGGVSVYKYFNSIRSCPRKFGQSHCEIDASKTCYHGNGDSYRKGANTDTKG 89
QY 69 RPLCPWNSATVLOQTTHAHRSDALQGLGKHCNCRPNRRPWCYVQVGLKPLVQECMV 128
DB 90 RCLAWNAVALQKPNNAHRSDAISLGLGKHCNCRPNQKRPWCYVQVGLRQFVQECMV 149
QY 129 HNCADGKFPSSPPEE 143
DB 150 HDCSLSKFPSSVDQ 164

RE JT 6
QY 0 PRELIMINARY; PRT; 214 AA.
AC Q9XT70;
DT 01-NOV-1999 (T-EMBLrel. 12, Created)
DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=lung;
RA Vin J., Idell S.;
RT Partial mRNA of rabbit uPA.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097647; AAD39351.1; -.
DR HSSP; P00749; LEJN.
DR MEROPS; S01.231; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR GlycoProtein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
FT NON_TER 214
QY SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match 40.3%; Score 337.5; DB 6; Length 214;
Best Local Similarity 78.7%; Pred. NO. 3.5e-28;
Matches 59; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

QY 67 MGRPCLPWSATVLOQTTHAHRSDALQGLGKHCNCRPNRRPWCYVQVGLKPLVQEC 126
DB 1 MDRPCLAWSANVLTQTHAHRSDALQGLGKHCNCRPNRRPWCYVQVGLKPLVQEC 60

QY 127 MVHDCADGKFPSSPP 141
DB 61 KVHD-SSGKKPALPP 74

```

```

RESULT 7
QY 727N2 PRELIMINARY; PRT; 291 AA.
ID Q727N2;
AC Q727N2;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Plasminogen activator, tissue type isoform 2.
GN PLAI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., da Ponte S.H., Hastings N.C.,
RA Alearn M.O., Kuldanek S.A., Rajkumar N., Toth E.J., Yi Q.,
RA Nickerson D.A.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY291060; AAP34246.1; -.
SQ SEQUENCE 291 AA; 32191 MW; 874E38C52F50EF1D CRC64;

Query Match 39.2%; Score 328.5; DB 4; Length 291;
Best Local Similarity 46.3%; Pred. NO. 4.5e-27;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP--SNCD---CLNGGTCVSNKYFNSHWCNCPKFGGQHCEIDSKTCYEGNGHGY 58
DB 77 QCHSVPVKSCSEPRFCNGGTCQALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 135
QY 59 RKGASTDTMRPCLPWSATVLOQTTHAHRSDALQGLGKHCNCRPNRRPWCYVQVGL 118
DB 136 RGTWSTAESGAECTNNWSSALAKQPSYGRPRDAIRLGLGNHNYCRNPRDRSKPWCYV 195
QY 119 LKPLVQECVHDCADG 134
DB 196 GKYSSSEFCSTPACSEG 211

RESULT 8
QY 86YK8 PRELIMINARY; PRT; 562 AA.
ID Q86YK8;
AC Q86YK8;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Tissue plasminogen activator.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RA Liu Y., Xu L., Zeng Y., He X.;
RT "cDNA of tissue plasminogen activator.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY221101; AA034406.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.

```


DR PROSITE; PS00021; KRINGLE 1; 2.
 DR PROSITE; PS00070; KRINGLE 2; 2.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 562 AA; 63668 MW; F956B4C77CB101E8 CRC64;

Query Match 37.2%; Score 311.5; DB 6; Length 562;
 Best Local Similarity 43.5%; Pred. No. 6.2e-25;
 Matches 60; Conservative 15; Mismatches 58; Indels 5; Gaps 3;

QY 3 ELHQP-SNCD---CLNGGTCVSNKYFNHWCNCKFKFGQCHCEIDKSKTCYEGNGHFY 58
 DB 77 QCHSVPVKSCSPRCFNGGTCLQALYFSD-F-VQCPVGVFGRCQCEIDARATCYEDQGIT 135
 QY 59 RGKASTDTMGRCPLPNSATVQLQTYHAHRSDALQLGLGKHNKNCNPNRRRPPWCYV 118
 DB 136 RGTWSTTESGAEVCVNWNTSGLASMPYNGRRPDAVKLGLGNHNYCRNPDKDSKPCWCIYFKA 195
 QY 119 LXPLVQECNVHDCADGKK 136
 DB 196 EKYSDFCSTPACTKEKE 213

RE: T 11
 Q: 11
 ID Q8K02 PRELIMINARY; PRT; 564 AA.
 AC Q8K02; 2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Tissue-type plasminogen activator.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBSJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DR EMBL; AY029518; AAK40240.1; -;
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibronectin.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR SMART; SM000395; Kringle; 2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; TYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS000395; Kringle; 2.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; TYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN; 1.
 DR PROSITE; PS00021; KRINGLE 1; 2.
 DR PROSITE; PS00070; KRINGLE 2; 2.

DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 564 AA; 62726 MW; 459D8BAC6D4A937C CRC64;

Query Match 36.0%; Score 301.5; DB 6; Length 564;
 Best Local Similarity 48.7%; Pred. No. 7.3e-24;
 Matches 56; Conservative 13; Mismatches 41; Indels 5; Gaps 3;

QY 5 HQVP-SNCD---CLNGGTCVSNKYFNHWCNCKFKFGQCHCEIDKSKTCYEGNGHFY 60
 DB 80 HSPVQVSCSPRCFNGGTCLQALYFSD-F-VQCPGVGRCRCEVDTRARCYEDRGIGY 138
 QY 61 KASTDTMGRCPLPNSATVQLQTYHAHRSDALQLGLGKHNKNCNPNRRRPPWCYV 115
 DB 139 TWTSTTESGACVNWNTSGLASMPYNGRRPDAVKLGLGNHNYCRNPDRTKPCWCIY 193

RESULT 12
 Q8K02
 ID Q8K02 PRELIMINARY; PRT; 517 AA.
 AC Q8K02; 2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBSJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; BC031775; AAH31775.1; -;
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 3.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
 KW Kringle; Protease; Serine protease.
 SQ SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;

```

Query Match          31.7%; Score 265; DB 11; Length 517;
Best Local Similarity 41.1%; Pred. No. 5.4e-20;
atches 53; Conservative 19; Mismatches 51; Indels 6; Gaps 4;

C 13 CLNGGTCVSNKYFSNIHWCNCPKKGQHCCEIDKSKTCYEGNGHYPYRGKASTDTMGRPCL 72
D 116 CQGGVCSRHRRSRF--TCACPDQYKGFCEIGPD--DCYVGDSYGRGKSVKTVNQNPCL 173
C 73 PWSNATVLOQTYHAHRSDALQLGLGHNYCRPNRRRWCVQVGLKPLVQE--CMVHDC 131
D 174 YNNSHLLLOETYNMFEDAEITHGIEHNFERNPDGDHRCWPCFKVNSKYKWEYCDVTVC 233
C 132 ADGKPPSP 140
D 234 ---PVDPDP 239

F ULT 13
C 520 PRELIMINARY; PRT; 560 AA.
A Q14520 O0663;
D 01-NOV-1996 (TREMELrel. 01, Created)
D 01-NOV-1996 (TREMELrel. 01, Last sequence update)
D 01-OCT-2003 (TREMELrel. 25, Last annotation update)
D HGF activator like protein (Hyaluronan binding protein 2).
C Homo sapiens (Human).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
C NCBI_TaxID=9606;
F (1)
K SEQUENCE FROM N.A.
F Kitamura N.;
F Submitted (MAR-1995) to the EMBL/GenBank/DBSJ databases.
F (2)
F SEQUENCE FROM N.A.
F MEDLINE=96425001; PubMed=8827452;
F Choi-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
F Tomita M.;
F "Purification and characterization of a novel hyaluronan-binding
F protein (PHEP) from human plasma: it has three EGF, a kringle and a
F serine protease domain, similar to hepatocyte growth factor
F activator.";
F J. Biochem. 119:1157-1165(1996).
F (3)
F SEQUENCE FROM N.A.
F TISSUE=Colon; and Kidney;
F Strausberg R.;
F Submitted (JUN-2002) to the EMBL/GenBank/DBSJ databases.
C -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
D EMBL; D49742; BAA08576.1; -.
D EMBL; S83182; AAB46909.1; -.
D EMBL; BC031412; AAH31412.1; -.
D FIR; JC4795; JC4795.
D HSP; P00763; IDPO.
D MEROPS; S01.033; -.
D Genew; HGNC:4798; HABP2.
D GO; GO:0005615; C:extracellular space; TAS.
D GO; GO:0005539; F:glycosaminoglycan binding; TAS.
D GO; GO:0007155; P:cell adhesion; TAS.
D InterPro; IPR009003; Cys_Ser_trypsin.
D InterPro; IPR006209; EGF-like.
D InterPro; IPR000001; Kringle.
D InterPro; IPR001254; Peptidase_S1.
D InterPro; IPR001314; Peptidase_S1A.
D Pfam; PF00008; EGF; 3.
D Pfam; PF00051; kringle; 1.
D Pfam; PF00089; trypsin; 1.
D PRINTS; PR00722; CHYMOTRYPSIN.
D PRINTS; PR00018; KRINGLE.
D ProDom; PD000395; Kringle; 1.
D PROSITE; PS00022; EGF 1; 3.
D PROSITE; PS01186; EGF_2; 2.

```

```

DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
SQ Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;

Query Match          31.5%; Score 263.5; DB 4; Length 560;
Best Local Similarity 41.0%; Pred. No. 8.6e-20;
Matches 55; Conservative 18; Mismatches 56; Indels 5; Gaps 4;

Qy 13 CLNGGTCVSNKYFSNIHWCNCPKKGQHCCEIDKSKTCYEGNGHYPYRGKASTDTMGRPCL 72
Db 159 CQGGATCSRHKRSKF--TCACPDQYKGFCEIG--SDCYVGDSYGRGKSVKTVNQHACL 216
Qy 73 PWSNATVLOQTYHAHRSDALQLGLGHNYCRPNRRRWCVQVGLKPLVQE--CMVHDC 131
Db 217 YNNSHLLLOETYNMFEDAEITHGIEHNFERNPDADAEKPCFIKVTNDKYKWEYCDVSAC 276
Qy 132 A--DGKPPSPPEE 143
Db 277 SAQDVAYPESPTPE 290

RESULT 14
O97507 PRELIMINARY; PRT; 616 AA.
AC O97507;
DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE FXII.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9623;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Takahashi T., Kihara T.;
RT "Porcine liver factor XII.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBSJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AB022426; BAA37148.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.211; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.

```

DR	SMART; SM00181; EGF; 2.
DR	SMART; SM00059; FN1; 1.
DR	SMART; SM00059; FN2; 1.
DR	SMART; SM00130; KR; 1.
DR	SMART; SM001020; TRYD_SPC; 1.
DR	PROSITE; PS00022; EGF_1; 2.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.
DR	PROSITE; PS00023; FIBRONECTIN_2; 1.
DR	PROSITE; PS00021; KINGLE_1; 1.
DR	PROSITE; PS00070; KINGLE_2; 1.
DR	PROSITE; PS00240; TRYPsin_DOM; 1.
DR	PROSITE; PS00134; TRYPsin_HIS; 1.
DR	PROSITE; PS00135; TRYPsin_SER; 1.
KW	EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW	Serine protease.
SC	SEQUENCE 616 AA; 68012 MW; 4CSFE3D7IEBBD1A9 CRC64;
	try Match 30.8%; Score 258; DB 6; Length 616;
	% Local Similarity 39.0%; Pred No. 3.7e-19;
	% Cons 57; Conservative 15; Mismatches 58; Indels 16; Gaps 5
Qy	6 QVPSNCDCINGGTCVSNKFNHWCNPKFGQGHCIDSKTCEYGEHGHPYRGASTD 65
Dd	176 QVCSTNPCLNGSGSLQTE---GHRLCRCPGTAGRLCDVDLKCRCYSRGLSYRGMAQT 232
Qy	66 TMGRPCLPNWATVLQQTY-HAHRSDALQLGLGKNYCENPDNRPRPPCYQVGKLPLVQ 124
Dd	233 LSGAPCQPWAS-----EATYNWTAEQALNWGLGDHFACRNPDNTRPWCVFWRGDLSMQ 288
Qy	125 ECMVHDCAD--GKKF-----SSPPE 142
Dd	289 YCLRARCQAPIGEAPPILTPQSPSE 314
RE	PRELIMINARY; PRT; 540 AA.
ID	Q800Y7
DT	01-JUN-2003 (TREMBLrel. 24, Created)
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE	Hepatocyte growth factor activator (Fragment).
OS	Meleagris gallopavo (Common turkey).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.
OX	NCBI_taxid=9103;
RN	[1]
RF	SEQUENCE FROM N.A.
RX	MEDLINE=2212796; PubMed=12128063;
RA	Holsberger D.R., Becker A.B., Thurston R.J., Rice C.D.;
RL	Submitted (JAN-2003) to the EMBL/GenBank/DDBSJ databases.
RL	EMBL; AY216598; AAC46038.1; -
DR	GO; GO:0005576; C:extracellular; IEA.
DR	GO; GO:0004263; F:cyclo-oxygenase activity; IEA.
DR	GO; GO:0004295; F:trypsin activity; IEA.
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR	InterPro; IPR009003; Cys_Ser_trypsin.
DR	InterPro; IPR000742; EGF 2.
DR	InterPro; IPR006209; EGF like.
DR	InterPro; IPR000083; Fibinctn1.
DR	InterPro; IPR000562; FN Type II.
DR	InterPro; IPR006210; IEGF.
DR	InterPro; IPR000001; Kringle.
DR	InterPro; IPR001254; Peptidase S1.
DR	InterPro; IPR001314; Peptidase SLA.
DR	Pfam; PF00008; EGF; 2.

DR	Pfam; PF00051; kringle; 1.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	PRINTS; PR00018; KRINGLE.
DR	ProDom; PD000395; Kringle; 1.
DR	SMART; SMART0181; EGF_2.
DR	SMART; SMART0058; FN1; 1.
DR	SMART; SMART00059; FN2; 1.
DR	SMART; SMART00130; KR; 1.
DR	SMART; SMART00020; Tryp_Spc; 1.
DR	PROSITE; PS00022; EGF_1; 2.
DR	PROSITE; PS01186; EGF-2; 1.
DR	PROSITE; PS01293; FIBRONECTIN_1; 1.
DR	PROSITE; PS00021; KRINGLE_1; 1.
DR	PROSITE; PSS0070; KRINGLE_2; 1.
DR	PROSITE; PSS0240; TRYPSIN_DOM; 1.
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.
DR	PROSITE; PS00135; TRYPSIN_SER; 1.
FT	NON TER 1 540 540
ET	NON TER
SQ	SEQUENCE 540 AA; 61159 MW; OBB3B4A89C0B577F CRC64;
Query Match	
Best Local Similarity 30.5%; Score 255.5; DB 13; Length 540;	
Matches	51; Conservative 17; Mismatches 54; Indels 11; Gaps 3;
Qy	13 CLNGGTC-----VSNKYFSNIHWNCNPKKFGCGHCEIDKSKTVEGNKHGYRGKASTDITMG 68
Dd	143 CMAGGECKMTASGGKV-----CDCKGPFFVKYCNIVPNHHCHVRNGTEYRGTAQTIGS 197
Qy	69 RPLPWNASATVLQOOTYHAHRSDALQLGLGHXCRNPENRRNRRCWCVVQVGKLPELVQECMV 128
Dd	198 HSCLPWNSDLYLREHLVDSEKAVQLGLGFPFCRNPFDEDEKEWCIMKDNSLSWEYCNI 257
Qy	129 HDCADGKKSPSP 141
Dd	258 TSCASRER--RPP 268
RESULT 16	
O8VC54	PRELIMINARY; PRT; 653 AA.
ID O8VC54 AC Q8VC54	
DT DT 01-MAR-2002 (TrEMBLrel. 20, Created)	
DD DD 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)	
DE DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
OS Hypothetical protein.	
OC Mus musculus (Mouse).	
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. [1]	
RN NCBI_Taxid=10090;	
RV SEQUENCE FROM N.A.	
RP RP TISSUE=Liver;	
RA Strauberg R.;	
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.	
CC -! SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.	
CC -! SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.	
DR HSSP; P00761; IAN1.	
DR GO; GO:0005576; C.extracellular; IEA.	
DR GO; GO:0004263; F.chymotrypsin activity; IEA.	
DR GO; GO:0008433; F.peptidase activity; IEA.	
DR GO; GO:0004295; F.trypsin activity; IEA.	
DR GO; GO:0006508; P.proteolysis and peptidolysis; IEA.	
DR InterPro; IPRO09003; Cys Ser tryptic.	
DR InterPro; IPRO00742; EGF-2.	
DR InterPro; IPRO06209; EGF-like.	
DR InterPro; IPRO00083; Fibnctnl.	
DR InterPro; IPRO00562; FN_type_I1.	
DR InterPro; IPRO06210; IEGF.	
DR InterPro; IPRO00001; Kringle.	
DR InterPro; IPRO01254; Peptidase S1.	

D: InterPro; IPR001314; Peptidase_S1A.
D: Pfam; PF00008; EGF; 2.
D: Pfam; PF00039; fn1; 1.
D: Pfam; PF00040; fn2; 1.
D: Pfam; PF00051; kringle; 1.
D: Pfam; PF00089; trypsin; 1.
D: PRINTS; PR00722; CHYMOTRYPSIN.
D: PRINTS; PR00013; FNTYPEII.
D: PRINTS; PR00018; KRINGLE.
D: ProDom; PD000995; FN Type II; 1.
D: ProDom; PD000395; Kringle; 1.
D: SMART; SM00181; EGF; 2.
D: SMART; SM00059; FN2; 1.
D: SMART; SM00130; KR; 1.
D: SMART; SM00020; Tryp SPC; 1.
D: PROSITE; PS00022; EGF 1; 2.
D: PROSITE; PS01186; EGF 2; 1.
D: PROSITE; PS01253; FIBRONECTIN_1; 1.
D: PROSITE; PS00023; FIBRONECTIN_2; 1.
D: PROSITE; PS00021; KRINGLE_1; 1.
D: PROSITE; PS00070; KRINGLE_2; 1.
D: PROSITE; PS00240; TRYPSIN_DOM; 1.
D: PROSITE; PS00134; TRYPSIN_HIS; 1.
D: PROSITE; PS00135; TRYPSIN_SER; 1.
K: Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
K: Kringle; Protease; Serine protease.
S: SEQUENCE 653 AA; 70553 MW; F18D90174ED6FDD CRC64;

Query Match 29.9%; Score 250.5; DB 11; Length 653;
Best Local Similarity 42.0%; Pred. No. 2.5e-18;
atches 50; Conservative 11; Mismatches 41; Indels 17; Gaps 2;

Q: 5 HQVSNCDLNGGTCVSNKYFSNTHW-----CNCPKFGGHCIDKSKTCYEGNGH 56
D: 239 HTACUSSPCLNGGTC-----HLVGTGTSVCTPLGVAGRCNIVPTEHFLNGT 289

Q: 57 FYRGKASTDTWGRPLCPWNSATVLQOYTHAHRSDALQLGLGKHNCRPNRRPWCYV 115
D: 290 EYRGVASTAAGSLSCLAWSNLLYQELHVDGVAALVLLGLGPHAYCRNPKDERPWCYV 348

R: ILT 17
Q: 27 O35727 PRELIMINARY; PRT; 597 AA.
A: O35727;
D: 01-JAN-1998 (Tremblrel. 05, Created)
D: 01-JAN-1998 (Tremblrel. 05, Last sequence update)
D: 01-OCT-2003 (Tremblrel. 25, Last annotation update)
D: Factor XII.
G: F12.
O: Mus musculus (Mouse).
O: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
O: Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
O: NCBI_TaxID=10090;
R: [1]
R: SEQUENCE FROM N.A.
R: TISSUE=Liver.
R: Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
R: Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
C: -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
C: -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
D: EMBL; X9571; CAA67891.1; -.
D: HSRP; P00760; 1AQ7.
D: MEROPS; S01.211; -.
D: MGD; MGI:1891012; F12.
D: GO; GO:0005576; C:extracellular; IEA.
D: GO; GO:0004263; F:chymotrypsin activity; IEA.
D: GO; GO:0008233; F:peptidase activity; IEA.
D: GO; GO:0004295; F:trypsin activity; IEA.
D: GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
D: InterPro; IPR000903; Cys_Ser_trypsin.
D: InterPro; IPR006209; EGF-like.
D: InterPro; IPR000083; Fibrinectn.

DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN Type II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp SPC; 1.
DR PROSITE; PS00022; EGF 1; 2.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;

Query Match 29.9%; Score 250; DB 11; Length 597;
Best Local Similarity 36.6%; Pred. No. 2.6e-18;
Matches 52; Conservative 16; Mismatches 54; Indels 20; Gaps 4;

QY 13 CLNGGFCVSNKYFSNTHWNCNPKFGGHCIDKSKTCYEGNGHFGYRGKASTDTWGRPCL 72
DB 183 CLNGGSCL---LVEDHPLCRCTGYTGVCDDLWATCYEGRLGYRGAGTQSGAPCQ 239

QY 73 PWSATVLTQOYTHAHRSDALQLGLGKHNCRPNRRPWCYVQVGLKPLVQECMVHDC 131
DB 240 RW----TVEATYRNMTKEQALSWGLGHAFCRNPNDTRPWCYVWSGDRLSWDYCGLEQC 295

QY 132 -----ADGKPKSSPP 141
DB 296 QTPTFAPLVVPESQESPSQAP 317

RESULT 18
ID Q80YCS PRELIMINARY; PRT; 609 AA.
AC Q80YCS;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Similar to coagulation factor XII (Hageman factor) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC049867; AAH49867.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.


```

D InterPro; IPR000001; Kringle.
D InterPro; IPR001254; Peptidase_S1.
D InterPro; IPR001314; Peptidase_S1A.
D Pfam; PF00008; EGF_2.
D Pfam; PF00039; fn1; 1.
D Pfam; PF00040; fn2; 1.
D Pfam; PF00051; kringle; 1.
D Pfam; PF00089; trypsin; 1.
D PRINTS; PR00722; CHYMOTRYPSIN.
D PRINTS; PR00013; FNTYPEII.
D PRINTS; PR00018; KRINGLE.
D ProDom; PD000995; FN Type II; 1.
D ProDom; PD000395; Kringle; 1.
D SMART; SM00181; EGF_2.
D SMART; SM00179; EGF_CA; 2.
D SMART; SM00058; FN1; 1.
D SMART; SM00059; FN2; 1.
D SMART; SM00130; KR; 1.
D SMART; SM00020; Tryp SPC; 1.
D PROSITE; PS00023; EGF_1; 2.
D PROSITE; PS01186; EGF_2; 1.
D PROSITE; PS01233; FIBRONECTIN_1; 1.
D PROSITE; PS00023; FIBRONECTIN_2; 1.
D PROSITE; PS00021; KRINGLE_1; 1.
D PROSITE; PS50070; KRINGLE_2; 1.
D PROSITE; PS50240; TRYPSIN_DOM; 1.
D PROSITE; PS00134; TRYPSIN_HIS; 1.
D PROSITE; PS00135; TRYPSIN_SER; 1.
D SEQUENCE 615 AA; 67735 MW; 030508970A0C7EDB CRC64;

ery Match
st Local Similarity 27.8%; Score 233; DB 4; Length 615;
atches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;

Q 13 CLINGGVSNKYSNHWNCNCPKKGQCEIDKSKTYEGNGHYRGKASTDTMGRPCL 72
D 183 CLINGGRCLE---VEGRLCHCPVGYTGPDCDDVTAKSCYDGRGLSYRLATITLGGAPCQ 239

Q 73 PWSATVLQOITY-HAHRSDALQLGLGKHNYCRNPDRRRPCWYC 115
D 240 PWAS----EATYENVTAEQARNWGLGGHAFACFPNDIRPCFV 279

R ILT 21
Q -20 PRELIMINARY; PRT; 90 AA.
I A QNG20
D 01-OCT-2002 (TREMBlrel. 22, Created)
D 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
D 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
D Plasminogen/activator kringle.
D Homo sapiens (Human).
D Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
D Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
D NCBI_TaxID=9606;
R [1]
R SEQUENCE FROM N.A.
R Dou D.;
R "Production of kringle fragment.";
R Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
R -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
D EMBL; AF282882; AA52248.1; --
D InterPro; IPR000001; Kringle.
D Pfam; PF00051; Kringle; 1.
D PRINTS; PR00018; KRINGLE.
D ProDom; PD000395; Kringle; 1.
D SMART; SM00130; KR; 1.
D PROSITE; PS00021; KRINGLE_1; 1.
D PROSITE; PS50070; KRINGLE_2; 1.
D Glycoprotein; Kringle.
R K SEQUENCE 90 AA; 9804 MW; A338879FDFA4C7B1 CRC64;

ery Match
st Local Similarity 26.1%; Score 218.5; DB 4; Length 90;
atches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;

Q 13 CLINGGVSNKYSNHWNCNCPKKGQCEIDKSKTYEGNGHYRGKASTDTMGRPCL 72
D 183 CLINGGRCLE---VEGRLCHCPVGYTGPDCDDVTAKSCYDGRGLSYRLATITLGGAPCQ 239

Q 73 PWSATVLQOITY-HAHRSDALQLGLGKHNYCRNPDRRRPCWYC 115
D 240 PWAS----EATYENVTAEQARNWGLGGHAFACFPNDIRPCFV 279

R ILT 21
Q -20 PRELIMINARY; PRT; 90 AA.
I A QNG20
D 01-OCT-2002 (TREMBlrel. 22, Created)
D 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
D 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
D Plasminogen/activator kringle.
D Homo sapiens (Human).
D Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
D Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
D NCBI_TaxID=9606;
R [1]
R SEQUENCE FROM N.A.
R Dou D.;
R "Production of kringle fragment.";
R Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
R -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
D EMBL; AF282882; AA52248.1; --
D InterPro; IPR000001; Kringle.
D Pfam; PF00051; Kringle; 1.
D PRINTS; PR00018; KRINGLE.
D ProDom; PD000395; Kringle; 1.
D SMART; SM00130; KR; 1.
D PROSITE; PS00021; KRINGLE_1; 1.
D PROSITE; PS50070; KRINGLE_2; 1.
D Glycoprotein; Kringle.
R K SEQUENCE 90 AA; 9804 MW; A338879FDFA4C7B1 CRC64;

Best Local Similarity 49.4%; Pred.No. 7.9e-16;
Matches 41; Conservative 5; Mismatches 36; Indels 1; Gaps 1;

QY 50 CYEGNGHYRGKASTDTMGRPCLPWNSATVLQOITYHAHRSDALQLGLGKHNYCRNPDRNR 109
Db 8 CYFGNGSAVRGTSHSLTESGASCLFWNSMLIKGVYTAQNPSAQALGLGKHNYCRNPDGDA 67
QY 110 RPWCIVQVGLKPLVQECMVHDCA 132
Db 68 RPPCYT-TNPKLYDYCDVPQCA 89

RESULT 22
Q90675 PRELIMINARY; PRT; 202 AA.
AC Q90675;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Tissue-type plasminogen activator (Fragment).
GN TPA.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn;
RX MEDLINE=97199025; PubMed=9047000;
RA Johnson A.L., Bridgham J.T., Anthony R.V.;
RT "Expression of avian urokinase and tissue-type plasminogen activator messenger ribonucleic acid during follicle development and atresia.";
RL Biol. Reprod. 56:581-588(1997).
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBU; U31988; AAA74955.1; --
DR HSSP; P00750; IRTF.
DR MEROPS; S01.232; --
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR SMART; SM00130; KR; 1.
DR ProDom; PD000395; Kringle; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 202
SQ SEQUENCE 202 AA; 23148 MW; 049DB42941D83AFC CRC64;

Query Match 23.6%; Score 197.5; DB 13; Length 202;
Best Local Similarity 36.9%; Pred.No. 3.4e-13;
Matches 41; Conservative 11; Mismatches 50; Indels 9; Gaps 2;

QY 30 WCNCPK--KFGQHCE-----IDSKTYEGNGHYRGKASTDTMGRPCLPWNSATVL 80
Db 12 WCYVFAGKYISEFCSTPACTKVAEEDGDCYTGNGLAYRGTRSRTRSGFSCLPWNPFVLT 71
QY 81 QCTYHAHRSDALQLGLGKHNYCRNPDRRRPCWYCVQVGLKPLVQECMVHD 131
Db 72 SKIYTALEEQRALGLGKHNYCRNPDRPDGAQPWCHWKDRQLTWECYCDVPOC 122

RESULT 23
Q9AXX3

```

```

Best Local Similarity 49.4%; Pred. No. 7.9e-16;
Matches 41; Conservative 5; Mismatches 36; Indels 1; Gaps 1;

QY 50 CYEGNGHYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNR 109
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8 CYFGNGSAYRGTHSLTSSGASCLFWSNMLIKGVYTAQNPASQAALGLGKHNYCRNPDGDA 67
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 110 RPWCYVQGLKPLVQECMVHDC 132
    :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 68 KPCYT-TNPKLYDYCDVPOCA 89
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 22
Q90675 PRELIMINARY; PRT; 202 AA.
AC Q90675;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Tissue-type plasminogen activator (Fragment).
GN TPA
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn;
RX MEDLINE=97199025; PubMed=9047000;
RA Johnson A.L., Bridgham J.T., Anthony R.V.;
RT "Expression of avian urokinase and tissue-type plasminogen activator
RT messenger ribonucleic acid during follicle development and atresia.";
RL Biol. Reprod. 56:581-588 (1997).
CC !- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; U31988; AAA74955.1; -.
DR HSSP; P00750; IRTF.
DR MEROPS; S01.232; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; trypsin; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00395; K-kringle; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin HIS; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 202
SQ SEQUENCE 202 AA; 23148 MW; 049DB42941D83AFC CRC64;

Query Match 23.6%; Score 197.5; DB 13; Length 202;
Best Local Similarity 36.9%; Pred. No. 3.4e-13;
Matches 41; Conservative 11; Mismatches 50; Indels 9; Gaps 2;

QY 30 WNCNPK-KFGQHC-----IDSKTYEGNGHYRGKASTDTMGRPCLPWNSATVL 80
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12 WCYVFKAGKYISFCSTPACTKVAEEDGDCYTGNGLAYRGTRSRKSGFSCLPWNVFLT 71
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 81 QCTYHAHRSDALQLGLGKHNYCRNPDNRFPWCYVQGLKPLVQECMVHDC 131
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 72 SKIYTALEEQRRALGLGKHNYCRNPDGDAQPWCHWVKDRQLTWEYCDVPOC 122
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 23
Q9AXX3

```


Q8AXX3 PRELIMINARY; PRT; 421 AA.
Q8AXX3
DT 01-WAR-2003 (TREMELrel. 23, Created)
DT 01-WAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Krenenz.
GN KRM2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22308873; PubMed=12421700;
RA Davidson G., Mao B., Del Barco Barrantes I., Niehrs C.;
RT "Kremen proteins interact with Dickkopf1 to regulate anteroposterior
RN CNS patterning."
RL Development 129:5587-5596(2002).
DR EMBL; AY150813; RAN64661.1; -.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00321; WSC; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
SQ SEQUENCE 421 AA; 46274 MW; 74D04DB1682CD15F CRC64;
Query Match 20.8%; Score 174; DB 13; Length 421;
Best Local Similarity 47.4%; Pred. No. 2.5e-10;
Matches 36; Conservative 6; Mismatches 28; Indels 6; Gaps 3;
QY 41 HCEIDSKTCYEGNGHFGYRGAS-DTWMGRPCLPWNSATVLQOTVHAHRSALQGLGKH 99
Db 23 HPELSE---CFTVNGRDVGTVSQAGPGTCLYWNQTT--QHLNAGSDPDGELGLGNH 77
QY 100 NYCRRPNDRRRPWCYV 115
Db 78 NYCRRPNDAVQPCVYV 93
RE IT 24
Q8AXX3 PRELIMINARY; PRT; 947 AA.
AC Q8AXX3
DT 01-WAR-2003 (TREMELrel. 23, Created)
DT 01-WAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Muscle-specific receptor tyrosine kinase Musk.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2038710; PubMed=11083926;
RA Ip F.C., Glass D.G., Gies D.R., Cheung J., Lai K.O., Fu A.K.,
RA Yancopoulos G.D., Ip N.Y.;
RT "Cloning and characterization of muscle-specific kinase in chicken."
RL Mol. Cell. Neurosci. 16:661-673(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Gies D., Glass D., Yancopoulos G.D.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY143173; AAN05008.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005874; C:microtubule; IEA.
DR GO; GO:0005824; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0007018; P:microtubule-based movement; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR002453; Beta_tubulin.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Sex_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008286; Tyr_kinase_AS.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; IG; 3.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 3.
DR SMART; SM00408; IGC2; 3.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00038; Fz; 1.
DR PROSITE; PS00835; IG-LIKE; 3.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS0107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00228; TUBULIN_B_AUTOREG; 1.
DR Receptor; Kinase.
SQ SEQUENCE 947 AA; 105588 MW; E6C3FCC5796CC8BE CRC64;
Query Match 20.8%; Score 174; DB 13; Length 947;
Best Local Similarity 31.1%; Pred. No. 6e-10;
Matches 50; Conservative 15; Mismatches 36; Indels 60; Gaps 10;
QY 4 LHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFGYRGAS 63
Db 441 LHQDPSAC-----THIPDFPKKE-----NITRTCYSGNGQFYGGWAN 478
QY 64 TDTMGRPCLPWNSATVLQOTVHAHR-----SDALQLGLGKKNYCRNP-DNRRRPMC 113
Db 479 VTASGIPCCQKWS-----DQAPHLHRRTPQVFFELSDA-----ENYCNPGGNERPMC 526
QY 114 YVQVGLKPLV--QECMVHDCAG-----KPPSS-----PP 141
Db 527 YTK---DPSVTWEYCVSVPCGDAS--SLGTRKPNGETQNLPP 564
RESULT 25
Q7SXB3
ID Q7SXB3 PRELIMINARY; PRT; 263 AA.
AC Q7SXB3
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

Cyprinidae; Danilo.
NCBI_TaxID=7955;
[1] SEQUENCE FROM N.A.
STRAIN=AB; TISSUE=Body;
MEDLINE=23388257; PubMed=12477932;
Strasberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Achsel S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan K.H., Moore T., Max S.I., Wang J., Heide F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Pahey J., Helton E., Kettner M., Madan A.C., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzyszewski M.J., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2] SEQUENCE FROM N.A.
STRAIN=AB; TISSUE=Body;
Strasberg R.;
Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
EMBL; BC055675; AAH55675.1; --
Hypothetical protein.
SEQUENCE 263 AA; 28777 MW; 8BEB1178C7C8A58 CRC64;

Query Match 19.8%; Score 166; DB 13; Length 263;
Best Local Similarity 35.7%; Pred. No. 1.1e-09;
Matches 35; Conservative 9; Mismatches 44; Indels 10; Gaps 3;

48 KTCYEGNGHFGYRGKASTDTMGRCPLPNSATVLOQTYHAHRSALQLGLGKHNYCRNPDN 107
23 KDCITNNGEDYRGTKQTSSTGSLWSRLNL-----KFKDSQTGVGDHNCRNPDG 74

108 RRRPWCYVQVGLKPLVQV-CMVHDCADGKPSPP-PEE 143
75 SNKPWCYVSGSGGKTKKACDIRICQDNATEAPAPEE 112

[1] PRELIMINARY; PRT; 716 AA.
Q91691
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Growth factor Livertine.
Xenopus laevis (African clawed frog).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus.
NCBI_TaxID=8355;
[1] SEQUENCE FROM N.A.
Ruiz i Altaba A., Thery C.;
Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
-!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
EMBL; U57455.1; AAB52574.1; --
HSP; P00747.1; ICAE.
GO; GO:0005509; F:calcium ion binding; IEA.
GO; GO:0004263; F:chymotrypsin activity; IEA.
GO; GO:0008233; F:peptidase activity; IEA.
GO; GO:0003809; F:thrombin activity; IEA.
GO; GO:0004295; F:trypsin activity; IEA.

DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR PRINTS; PRO1505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLS_1; 4.
DR PROSITE; PS0070; KRINGLS_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR GlycoProtein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 716 AA; 81971 MW; 508376A0E4398798 CRC64;

Query Match 19.8%; Score 166; DB 13; Length 716;
Best Local Similarity 31.1%; Pred. No. 3.2e-09;
Matches 42; Conservative 11; Mismatches 40; Indels 42; Gaps 6;

QY 30 WC-----NCPKFGGQHCIDKSK-----TCYEGNGHFGYRGKASTDTMGRCPLP 73
DB 258 WCYTTDPNVEKEF---CRITCKQRLSNILITSCPKRGEYRGKANTTTSIGIPCOR 313

QY 74 WNSAT-----VLQOTYHAHRSALQLGLGKHNYCRNPDNRRPWCYVQV-----GL 119
DB 314 WDSQTPQSHRFLPEKYPCKGLD-----ENYCRNPDGSEAPWCFTTLPGMRMAYCFQI 365

QY 120 KPLVQECMVHDCADG 134
DB 366 KRCKDDVLEPDCVHG 380

RESULT 27
Q8WMR1 ID Q8WMR1 PRELIMINARY; PRT; 359 AA.
AC Q8WMR1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1] SEQUENCE FROM N.A.
RP Parle-Shepherd S.R., Coffman K.T., Resnick D., Chan R., Kisker O.,
RA Folkman J., Waters D.J.;
RT "Angiostatin is Detectable in the Urine of Dogs with Spontaneous Bone
RT Cancer";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; AY069985; AAL5819.1; --
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 4.
DR PRINTS; PRO0018; KRINGLE.
DR PRINTS; PRO1505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.

```
DR SMART; SM00130; KR; 4.
DR PROSITE; PS00021; KRINGLE 1; 4.
DR PROSITE; PS00070; KRINGLE 2; 4.
DR KW Glycoprotein; Kringle.
DR FT NON_TER 1
DR FT NON_TER 359
DR FT NON_TER 359
DR SQ SEQUENCE 359 AA; 41172 MW; 776D35F4AB0BDD9E CRC64;
Query Match 19.8%; Score 165.5; DB 6; Length 359;
Best Local Similarity 29.0%; Pred.No. 1.7e-09;
Matches 42; Conservative 13; Mismatches 43; Indels 47; Gaps 7;
QY 5 HQVPSNCDCLNGTGVSNKYFSN-----IHWK-NCPKFGGQHCIDKSKT----- 49
Dd 211 NRTPEFPCKN-----LDENYCRNPDEGTAPWCYTTNSEVRWEHCIPSCSSPITTEYLD 266
QY 50 -----CYEGNGHFYRGKASTDTMGPRCLPWSATVLOQTYHAHRSAL-- 92
Dd 267 APASVPEPTPVQECYQNGKSYRGTSSTTTGKCKOSWSMT-----PHRHEKTP 319
QY 93 ---QLGLGHNYCRNPDRRRPWCY 114
Dd 320 HPPEAGL-TMNYCRNPDAKSPWCY 343
RE T 28
QY 43 Q9ROW3 PRELIMINARY; PRT; 812 AA.
AC Q9ROW3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
GN PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Liver;
RA Bangert K., Johnsen A.H., Thorsen S.;
RT "Rat plasminogen: cDNA and gene structure.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Tissue=Liver;
RA Kanalas J.J., Makker S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829(1991).
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; A242649; CAB46014.1; -.
DR HSP; P00747; IPMK.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004263; F:hormone activity; IEA.
DR GO; GO:0004283; F:peptidase activity; IEA.
DR GO; GO:0004283; F:plasmin activity; IEA.
DR GO; GO:0003809; F:fibrinogen activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR InterPro; IPR001400; Somatotropin.
```

```
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR0018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 5.
DR PROSITE; PS00070; KRINGLE 2; 5.
DR PROSITE; PS00338; SOMATOTROPIN 2; 1.
DR PROSITE; PS00240; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KW Glycoprotein; Hydroxylase; Kringle; Protease; Serine protease; Signal.
FT SIGNAL 1
FT CHAIN 19
FT CHAIN 20 812 PLASMINOGEN.
SQ SEQUENCE 812 AA; 90535 MW; 8C703C51410BC9E CRC64;
Query Match 19.8%; Score 165.5; DB 11; Length 812;
Best Local Similarity 27.8%; Pred.No. 4.2e-09;
Matches 50; Conservative 22; Mismatches 51; Indels 57; Gaps 12;
QY 5 HQVPSNCDCLNGTGVSNKYFSN-----IHWK-NCPKFGGQHCIEI-----DKS-- 47
Dd 308 NRTPEFPCKN-----LDENYCRNPDEGTAPWCYTTDSQLRWYCEIPSCGSSVSPQSDS 363
QY 48 -----KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLOQTYHAHRSALQL-- 94
Dd 364 SVLPETPVVQECYQNGKSYRGTSSTTTGKCKOSW-----VSMTPHSHSKTPANFPDA 418
QY 95 GLGHNYCRNPDRRRPWCYQVGLKPLV--QECMVHDCAD-----GKXPSSP 140
Dd 419 GL-EMNYCRNPDRNDQRPWCFT--TDPSVRWEYCNLKECSETGGGVAESAIVPQVPSAP 474
RESULT 29
Q9BRB6 PRELIMINARY; PRT; 393 AA.
AC Q9BRB6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to receptor tyrosine kinase-like orphan receptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Muscle;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC006374; AA06374.1; -.
DR HSP; P00747; ICEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
```

D1	PROSITE; PS00038; FZ; 1.	DR	SMART; SM00408; IGc2; 1.
D1	PROSITE; IG LIKE; 1.	DR	SMART; SM00130; KR; 1.
D1	PROSITE; PS00021; KRINGLE_1; 1.	DR	SMART; SM00220; S_TKC; 1.
D1	PROSITE; PS00070; KRINGLE_2; 1.	DR	SMART; SM00219; Tyrc; 1.
X1	Glycoprotein; Immunoglobulin domain; Kinase; Kringle; Receptor.	DR	PROSITE; PS00038; FZ; 1.
S1	SEQUENCE 393 AA; 43825 MW; 1F93DCBBBF3855 CRC64;	DR	PROSITE; PS00021; KRINGLE_1; 1.
		DR	PROSITE; PS00035; IG LIKE; 1.
		DR	PROSITE; PS00070; KRINGLE_2; 1.
		DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
		DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
		SQ	SEQUENCE 937 AA; 104118 MW; CB440323CA66EC28 CRC64;
			Query Match 18.9%; Score 158.5; DB 11; Length 937;
			Best Local Similarity 29.7%; Pred. No. 2.7e-08;
			Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;
			2 NELHQP-----SNDCINGTCVSNKYPS-----NIHCNC-----PKFGGQH 41
			236 DETSVKPRDLCDCEILEVLCQTEYIFARSPNMLMLKLPNCEDLPQSPESPAAN 295
			42 C-----EIDSKTCYEGNGHFYRGKASTDWMGRPCLPWSATVLQQTVAHRSAL 92
			296 CIRIGIPWADPINKHKCYNSTGVYRGTVSVTKSGRCQCPWNS-----QYPHTHSFTAL 350
			93 QLG--LGKHYCRNPDRRR--PCWY 114
			351 RPELNGGHSYCRNPGNQKEAPWCF 375
			QSEN99 PRELIMINARY; PRT; 937 AA.
			01-MAR-2003 (TREMBlrel. 23, Created)
			01-MAR-2003 (TREMBlrel. 23, Last sequence update)
			01-OCT-2003 (TREMBlrel. 25, Last annotation update)
			Receptor tyrosine kinase-like orphan receptor 1.
			RORL
			Mus musculus (Mouse).
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
			Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
			NCBI_TaxID=10090;
			[1]
			SEQUENCE FROM N.A.
			STRAIN=C57BL/6J; TISSUE=Adipose tissue;
			MEDLINE=22354683; PubMed=12466851;
			The FANTOM Consortium,
			"Analysis of the mouse transcriptome based on functional annotation of
			60,770 full-length cDNAs."
			Nature 420:563-573 (2002).
			EMBL; AK080896; BAC38068.1; --
			MGD; MGI:1347520; Rorl.
			GO; GO:0006488; P:protein amino acid phosphorylation; IEA.
			InterPro; IPR000024; Fz domain.
			InterPro; IPR003599; IG.
			InterPro; IPR007110; IG-like.
			InterPro; IPR003598; IG c2.
			InterPro; IPR000001; Kringle.
			InterPro; IPR000719; Prot_kinase.
			InterPro; IPR002290; Ser_thr_kinase.
			InterPro; IPR001245; Tyr_pkinase.
			InterPro; IPR008266; Tyr_pkinase_AS.
			Pfam; PF01392; Fz; 1.
			Pfam; PF00047; Ig; 1.
			Pfam; PF00051; kringle; 1.
			Pfam; PF00069; pkinase.1.
			PRINTS; PR00018; KRINGLS.
			PRINTS; PR00109; TYRKINASE.
			ProDom; PD000395; Kringle; 1.
			ProDom; PD000001; Prot_kinase; 1.
			SMART; SM00409; IG; 1.

DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PRO0018; KRINGLE.
 DR PRINTS; PRO0019; TYRKINASE.
 DR ProDom; PD000395; Kringle; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00220; S_TK; 1.
 DR SMART; SM00219; Tyrc; 1.
 DR PROSITE; PS00038; FZ; 1.
 DR PROSITE; PS50035; IG LIKE; 1.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 SQ SEQUENCE 937 AA; 10408 MW; D6F2D84567D03C69 CRC64;
 Query Match 18.9%; Score 158.5; DB 11; Length 937;
 Best Local Similarity 29.7%; Pred. No. 2.7e-08;
 Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;
 QY 2 NELHQP-----SNDCINGTCTVSNKYFS-----NIHWNC-----PKFGQGH 41
 DB 236 DETSVKPRDLRDECEVLENLVLCQTEYIFARSNPMLMELKLPNCEDLPQESPEAN 295
 QY 42 C-----BIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNSTVQLQTYHAHRSDAL 92
 DB 296 CIRGIPNADPINKHKYCNSTGVYRGTVTSKGRCCQPNWS-----QYPHTSHSTAL 350
 QY 93 QLG--LGKHNCRPNRR--PCWY 114
 DB 351 RPPELNGGHSYCRNPGNKEAPWCF 375
 RE 4 T 32
 ID 046506 PRELIMINARY; PRT; 454 AA.
 AC 046506;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Apolipoprotein a (Fragment).
 GN BABAPOA.
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Papio.
 OX NCBI_TaxID=9557;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cox L.A., Jett C., Hixson J.E.;
 RT Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice Site Mutation is Associated with Deletion of a Single Exon in a Null Allele.
 RT Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DR EMBL; AF029691; AAB97886.1; -.
 DR HSSP; P00747; 2PK4.
 DR MEROPS; S01.999; -.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PRO0722; CHYMOTRYPSIN.
 DR PRINTS; PRO0018; KRINGLE.

DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00021; KRINGLE 1; 2.
 DR PROSITE; PS00070; KRINGLE 2; 2.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;
 KW Serine protease.
 FT NON_TER 1
 SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;
 Query Match 18.9%; Score 158; DB 6; Length 454;
 Best Local Similarity 39.8%; Pred. No. 1.4e-08;
 Matches 35; Conservative 10; Mismatches 31; Indels 12; Gaps 4;
 QY 50 CYEGNGHFYRGKASTDTMGRPCLPWNSTVQLQ--QTYHAHRSDALQLGLGKHNCRPN 107
 DB 102 CYHGDQSYRGSTFTVTGRTCSWSMTPHQKRTPEHPNDGLTM----NYCRNDA 156
 QY 108 RRRPCYVQVGLKPLV--QECNVHDCAD 133
 DB 157 DTGFWCFT---MDPSVRWEYCNLTTCSD 181
 RESULT 33
 ID Q9UIR5 PRELIMINARY; PRT; 113 AA.
 AC Q9UIR5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Apolipoprotein(a) (Fragment).
 GN APOA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21181705; PubMed=11285247;
 RX Ogorekova M., Kraft H.G., Ehnholm C., Utermann G.;
 RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV types 6 to 10 domain affect Lp(a) plasma concentrations and have different patterns in Africans and Caucasians."
 RT Hum. Mol. Genet. 10:815-824(2001).
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; AF158663; AAF03680.1; -.
 DR EMBL; AF158662; AAF03680.1; JOINED.
 DR HSSP; P00747; 1PMK.
 DR InterPro; IPR000001; Kringle.
 DR Pfam; PF00051; kringle; 1.
 DR PRINTS; PRO0018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 1.
 KW Glycoprotein; Kringle; Lipoprotein.
 FT NON_TER 1
 FT NON_TER 113
 SQ SEQUENCE 113 AA; 12685 MW; F3D65681D9B5253A CRC64;
 Query Match 18.8%; Score 157; DB 4; Length 113;
 Best Local Similarity 34.8%; Pred. No. 4e-09;
 Matches 39; Conservative 13; Mismatches 36; Indels 24; Gaps 6;
 QY 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSTV--LQTYHAHRSDALQLGLGKHNCRNP 105
 DB 9 RCYHNGQSYRGSTFTVTGRTCSWSMTPHRQRTPEHPNDGLTM----NYCRNP 63
 QY 106 DNRERPWCYVQVGLKPLV--QECNVHDCADG-----KKPS-SPPEE 143
 DB 64 DADTGEWCFT---MDPSVRWEYCNLTTCSDTEGTVPVAPPTVIQVPSLGPPE 112

```

R1 ILT 34
Q 84 Q7TP84 PRELIMINARY; PRT; 759 AA.
AC Q7TP84;
D 01-OCT-2003 (TRENBLrel. 25, Created)
D 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
D 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
D Abi-346.
C Rattus norvegicus (Rat).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
C NCBI_TaxID=10116;
R1 SEQUENCE FROM N.A.
R1 Xu C.S., Li W.Q., Li Y.C., Yan H.M., Chang C.F., Zhao L.F., Ma H.,
R1 Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
R1 Yang K.J., Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
R1 "Liver regeneration after PH.";
R1 Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
D EMBL; AY325159; AAP2560.1; -.
S SEQUENCE 759 AA; 86056 MW; 1422BFAC05C6DFA7 CRC64;

Query Match 18.7%; Score 156.5; DB 11; Length 759;
Best Local Similarity 24.2%; Pred. No. 3.6e-08;
Matches 53; Conservative 22; Mismatches 49; Indels 95; Gaps 12;

Q 5 HQVPSNCDCLNGGTCVSNKYFSN-----IHW-NCPKFGGQHCIEI-----DKS-- 47
D 316 NRTPEPFCKN-----LEENYCRNPDGETAPWCYTDSQLRWEYCEIPSCGSSVSPDQSDS 371
Q 48 -----KTCYEGNGHFYRGKASTDTCGRCLPWNSATVLOQTYHAHS-----DA 91
D 372 SVLPQETPVQECYQNGKSYRGTSITNTGKKQSW-----VSMTHSKSKTPANFPD 426
Q 92 LQLGLGKH-----NYCRNPDN-RRREWY 114
D 427 YONLLIQHWRPSVYVMSLWETFLIHTNQYSVQLKSSGLEWNYCRNPDNDQGEWCF 486
Q 115 VQVGLKPLV--QECWVHDCAD-----GKKPSP 140
D 487 T--TDPFSVRWEYCNLKRCSETGGTGGVAESAIVPQVPSAP 522

R1 ILT 35
Q 46 Q15146 PRELIMINARY; PRT; 810 AA.
AC Q15146;
D 01-NOV-1996 (TRENBLrel. 01, Created)
D 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
D 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
D Plasmimogen precursor.
C Homo sapiens (Human).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
C NCBI_TaxID=9606;
R1 SEQUENCE FROM N.A.
R1 TISSUE=Liver;
R1 Browne M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
R1 Mitchell D., Robinson J.H.;
R1 "Expression of recombinant human plasminogen and aglycoplasminogen in
R1 HeLa cells.";
R1 Fibrinolysis:0-0-0(1991).
D -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
D EMBL; M74220; AAA36451.1; -.
D HSP; P00747; 2PK4.
D GO; GO:0005509; F:calcium ion binding; IEA.
D GO; GO:0004263; F:chymotrypsin activity; IEA.
D GO; GO:0008233; F:peptidase activity; IEA.
D GO; GO:0003809; F:thrombin activity; IEA.
D GO; GO:0004295; F:trypsin activity; IEA.

```

```

DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR PRINTS; PRO1505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRY_SP; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS00340; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 810 PLASMINOGEN.
SQ SEQUENCE 810 AA; 90555 MW; B05C7D4B0D020B3C CRC64;

Query Match 18.7%; Score 156.5; DB 4; Length 810;
Best Local Similarity 28.5%; Pred. No. 3.8e-08;
Matches 49; Conservative 22; Mismatches 52; Indels 49; Gaps 12;

QY 5 HQVPSNCDCLNGGTCVSNKYFSN-----IHWNCNP-KKFGGQHCIEI-----DKS----- 47
DB 308 NRTPEPFCKN-----LDENYCRNPDGKAPWCHTNTSQRWEYCKIPSCDSSPVSTEQLA 363
QY 48 -----KTCYEGNGHFYRGKASTDTCGRCLPWNSATVLOQTYHAHS-----SDALQ 93
DB 364 PTAPPPLTPVQDCYHGDSYRGTSITNTGKKQSWSS-----WTHRHQKTEYENFN 418
QY 94 LGLGKHNYCRNPDNRRRWYVQVGLKPLV--QECWVHDCADGKKPS--SPP 141
DB 419 AGL-TMNYCRNPDADKGPWCFT---TDPFSVRWEYCNLKKS-GTEASVVAPP 465

RESULT 36
Q8KQ08 PRELIMINARY; PRT; 801 AA.
AC Q8KQ08;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC030848; AAH30848.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004574; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

```


68 RPWCYT---MDPSVWEYCNLTRC 89

Db RE JT 41
P7: 16
ID P70006 PRELIMINARY; PRT; 717 AA.
AC P70006;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hepatocyte growth factor-like protein precursor.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=96404125; PubMed=8808403;
RA Alberger F., Schmidt G., Richter K.;
RT "The Xenopus homologue of hepatocyte growth factor-like protein is
RT specifically expressed in the presumptive neural plate during
RT gastrulation."
RL Mech. Dev. 54:23-37(1996).
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; Y08734; CAA69989.1; -.
DR HSSP; P00747; 1CEA.
DR MEROPS; S01.977;
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRYSP; 1.
DR PROSITE; PS00021; KRINGLE_1; 3.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR GlycoProtein; Hydrolase; Kringle; Protease; Serine protease; Signal.
FT SIGNAL 1 28 POTENTIAL
FT CHAIN 29 717 HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.
SQ SEQUENCE 717 AA; 82017 MW; 6F8F77A432C8CDD54 CRC64;

Query Match 18.2%; Score 152.5; DB 13; Length 717;
Best Local Similarity 31.7%; Pred. No. 9e-08;
Matches 39; Conservative 13; Mismatches 42; Indels 29; Gaps 5;

QY 30 WC-----NCPKFGGQHCEIDK-----TCYEGNGHYRGKASTDTMGRCLP 73
Db 259 WCYTDPNVEF-----CHITCKQRISNITSTCFKRGEGYRGKANTTSGIFCQR 314
QY 74 WNSATVLOQTYHAHSDALQL---GLGHNYCRPNDRRRPWCYVQVGLKPLVQECMVHD 130
Db 315 WDT-----QAPHVHFLEPKYCKGLDE-NYCRAPVGSSEAPWCFITLKNWEMAYCFQIKR 369

QY 131 CAD 133
Db 369 CTD 371

RESULT 42
Q13208 PRELIMINARY; PRT; 567 AA.
AC Q13208;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hepatocyte growth factor-like protein homolog.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20191171; PubMed=10728827;
RA Degen S.J.F., McDowell S.A., Wultz S.E., Gould F., Stuart L.A.,
RA Carritt B.;
RT "Structure of the human D1F15S1A locus: a chromosome 1 locus with 97%
RT identity to the chromosome 3 gene coding for hepatocyte growth factor-
RT like protein."
RL DNA Seq. 8:409-413(1998).
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; U28054; AAC83092.1; -.
DR HSSP; P00747; 2PK4.
DR MEROPS; S01.977; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00020; TRYSP; 1.
DR PROSITE; PS00021; KRINGLE_1; 3.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR GlycoProtein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 567 AA; 64117 MW; 3FC38807F1645810 CRC64;

Query Match 18.2%; Score 152; DB 4; Length 567;
Best Local Similarity 29.4%; Pred. No. 7.9e-08;
Matches 37; Conservative 17; Mismatches 46; Indels 26; Gaps 7;

QY 30 WC-NCPKFGGQHCEIDK-----SKTYEGNGHYRGKASTDTMGRCLPWN 76
Db 225 WCYTDPQIEREFCDLPCGSEAPQBEATSVSCFPGKGEYRGKANTTGTAGVPCQRWDA 284
QY 77 ATVLQOTYHAHSDALQLGLG---KINYCNPNDRRRPWCYVQVGLKP--LVQEC-WVHDC 131
Db 285 -----QIPQHRFTPEKYACKDLRENFCNPNPDGSEAPWCFT---LRPGRVGFYQIRRC 336
QY 132 ADGKKP 137
Db 337 TDDVRP 342

R JLT 43
Q I V4 PRELIMINARY; PRT; 648 AA.
A Q9HIV4;
D 01-MAR-2001 (TRENBLrel. 16, Created)
D 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
D 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
D DJ1182A14.3 (Similar to MST1 (Macrophage stimulating 1 (Hepatocyte growth factor-like)))
D DJ1182A14.3.
C Homo sapiens (Human).
O Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
O Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
C NCBI_TaxID=9606;
R [1]
R SEQUENCE FROM N.A.
R Bird C.;
R Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
R -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
C -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
D HMBP; ALI37798; CAC17639.1; -.
D HSP; P00747; SHPG.
D GO; GO:0004263; F:chymotrypsin activity; IEA.
D GO; GO:0008233; F:peptidase activity; IEA.
D GO; GO:0004295; F:trypsin activity; IEA.
D GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
D InterPro; IPR009003; CYP_Ser_trypsin.
D InterPro; IPR000001; Kringle.
D InterPro; IPR003014; PAN.
D InterPro; IPR003609; Pan app.
D InterPro; IPR001254; Peptidase S1.
D InterPro; IPR001314; Peptidase_S1A.
D Pfam; PF00024; PAN; 1.
D Pfam; PF00051; kringle; 4.
D Pfam; PF00089; trypsin; 1.
D PRINTS; PR00722; CHYMOTRYPSIN.
D PRINTS; PR00018; KRINGLE.
D ProDom; PD000395; Kringle; 4.
D SMART; SM00130; KR; 4.
D SMART; SM00473; PAN AP; 1.
D SMART; SM00020; TRYP_SPC; 1.
D PROSITE; PS00021; KRINGLE_1; 3.
D PROSITE; PS00070; KRINGLE_2; 4.
D PROSITE; PS00240; TRYPSIN_DOM; 1.
D Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
S SEQUENCE 648 AA; 72781 MW; 4C80770573508463 CRC64;

Query Match 18.2%; Score 152; DB 4; Length 648;
Best Local Similarity 29.4%; Pred. No. 9.1e-08;
Matches 37; Conservative 17; Mismatches 46; Indels 26; Gaps 7;

Q 30 WC-NCPKFGGQHCIDK-----SKTCYEGNGHFYRGKASTDTMGRPCLPWS 76
D 205 WCYTTDPQIEREFCDLPRGSEAQPROBATSVCFRGKGEGYGTANTTAGVFCORWDA 264
Q 77 ATVLQOTYHAHSDALQLGLG--KNYCNPNRRRRCYQVGLKP--LVQEC-MVHDC 131
D 265 -----QIPHOHRTPEKYACKDLRENFCEPNPDGSEAPWCFT---LRPGRVGFCYQIERC 316
Q 132 ADGKKP 137
D 317 TDDVRP 322

R JLT 44
Q I V4 PRELIMINARY; PRT; 930 AA.
A Q8AV69;
D 01-MAR-2003 (TRENBLrel. 23, Created)
D 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
D 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
C Receptor tyrosine kinase Xror2.
C XRO2.

OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=2286220; PubMed=12399314;
RA Hikasa H., Shibata M., Hiratani I., Taira M.;
RT "The Xenopus receptor tyrosine kinase Xror2 modulates morphogenetic movements of the axial mesoderm and neuroectoderm via Wnt signalling.";
RL Development 129:5227-5239(2002).
DR EMBL; AB087137; BAC16209.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000024; Pz_domain.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF01392; Pz; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TyTK; 1.
DR PROSITE; PS00038; Pz; 1.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Kinase.
SQ SEQUENCE 930 AA; 104081 MW; C68454572411A8B6 CRC64;

Query Match 18.0%; Score 150.5; DB 13; Length 930;
Best Local Similarity 28.0%; Pred. No. 2e-07;
Matches 37; Conservative 18; Mismatches 44; Indels 33; Gaps 8;

Qy 11 CDCNGTCTVSNKY-----FSNIHMCNC-----PKFGGQHC-----EIDSK 48
D 252 CEVLNDLC-RQSYNTARSNPLILMQLHLPNCBELPLPESPEANCMRIGIPVEKLNRYQ 310
Qy 49 TCYEGNGHFYRGKASTDTMGRPCLPWSNATVLOOTYHAH---RSDALQLGLGHNYCRNP 105
D 311 QCYNGTGTDRGVSVTKGHCQCPWS-----HQVPHSHLSNADYPTG-GGHSYCRNP 364
Qy 106 DNRRL-PWCYVQ 116
D 365 GGQMEGWCFQ 376

RESULT 45
Q7ZTN9 PRELIMINARY; PRT; 709 AA.
ID Q7ZTN9
AC Q7ZTN9;

DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Similar to macrophage stimulating 1 (Hepatocyte growth factor-like).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RF SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.; Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL, BC044008; AAH4008.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR01254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR001966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE 1; 4.
DR PROSITE; PS00070; KRINGLE 2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
SQ SEQUENCE 709 AA; 81123 MW; 51C44D8EBD03B76A CRC64;
Query Match 17.9%; Score 149.5; DB 13; Length 709;
Best Local Similarity 28.0%; Pred. No. 1.9e-07;
Matches 44; Conservative 16; Mismatches 58; Indels 39; Gaps 8;
Qy 15 NGGTC-----VSNKYFSGNIHW-----CRCP-----KKFGGQHCEIDKS 47
Db 127 NGRTCHWLKFPDHKFSPTWPELENYCRNFDSDPEGLWCYTDTKNIHQYCGIKK 186
Qy 48 K--TCVEGCHFYRGKASTDTMGPRCLPWNSATVLOQTY-HAHRSDALQLGLGKKNYCRN 104
Db 187 EDVAVLCITCGEDYRGVSDRTSGKSCQRWD---LQAPHTFPYKPKYKPKDLSDDNYCRN 242
Qy 105 PDNRPRWCYVQVGLKPLVQ--CMVHDCADGKKPSS 139
Db 243 PDSSRPWCYT---TDPNVEREFCHITKIEKORISN 276
REUT 46
Qy 8
ID Q3UIR8 PRELIMINARY; PRT; 105 AA.
AC Q3UIR8;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RF SEQUENCE FROM N.A.
RC MEDLINE=21181705; PubMed=11285247;
RA Ogorelkova M., Kraft H.G., Ehnholm C., Utermann G.; "Single nucleotide polymorphisms in exons of the apo(a) kringles IV types 6 to 10 domain affect Lp(a) plasma concentrations and have different patterns in Africans and Caucasians."; Hum. Mol. Genet. 10:815-824(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR ENBL; AF158656; AAF03677.1; -.
DR HSSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 105
FT NON_TER 105
SQ SEQUENCE 105 AA; 11892 MW; 6ECB6C02CD30EFA2 CRC64;
Query Match 17.7%; Score 148; DB 4; Length 105;
Best Local Similarity 38.1%; Pred. No. 3.4e-08;
Matches 32; Conservative 8; Mismatches 36; Indels 8; Gaps 3;
Qy 50 CYEGNCHFYRGKASTDTMGPRCLPWNSATVLOQTYHAHRSDALQLGLGKKNYCRNPNRR 109
Db 11 CYHGDGSGYRGFSFTVTGRTCSWSMT---PHHQRTTEYYPNGGLTNYCRNPDAAI 67
Qy 110 RPMCIVQVGLKPLV--QECMVHDC 131
Db 68 SPWCYT---MDPNVWEYCNLTQC 88
RESULT 47
ID Q90865 PRELIMINARY; PRT; 704 AA.
AC Q90865;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hepatocyte growth factor-like/macrophage stimulating protein.
GN HGFI/MSP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RF SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=96029010; PubMed=7554499;
RA Thery C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.; "Expression of HGF/SF, HGFI/MSP and c-met suggests new functions during early chick development."; Dev. Genet. 17:90-101(1995).
RL Dev. Genet. 17:90-101(1995).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR ENBL; X84043; CAA58862.1; -.
DR HSSP; P00747; 1CEA.
DR MEROPS; S01.977; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.

```
D: InterPro; IPR001254; Peptidase S1.
D: InterPro; IPR001314; Peptidase_S1A.
D: Pfam; PF00051; kringle_4.
D: Pfam; PF00024; PAN; 1.
D: Pfam; PF00089; trypsin; 1.
D: PRINTS; PR00722; CHYMOTRYPSIN.
D: PRINTS; PR00018; KRINGLE.
D: ProDom; PD000395; Kringle; 4.
D: SMART; SM00130; KR; 4.
D: SMART; SM00473; PAN_AP; 1.
D: SMART; SM00020; TRYD_SPC; 1.
D: PROSITE; PS00021; KRINGLE_1; 4.
D: PROSITE; PS50070; KRINGLE_2; 4.
D: PROSITE; PS50240; TRYPSIN_DOM; 1.
D: PROSITE; PS50240; TRYPSIN_DOM; 1.
K: Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
S: SEQUENCE 704 AA; 79341 MW; CAB0D8CC41367C37 CRC64;

Query Match 17.7%; Score 148; DB 13; Length 704;
Best Local Similarity 30.9%; Pred. No. 2.7e-07;
Matches 47; Conservative 12; Mismatches 63; Indels 30; Gaps 10;

Q 5 HQVSNDC--LNGTGVSNKYFSNIHWCNCPKX-----FGQHC--RIDSKTCYEG 53
D: 312 HFVPSKYPCKDLQENYC--RNDGSEAPWCFTTREGMRVAFCHIRCDDEL--ABECYHG 369
Q 54 NGHYRGKASDTMTGRPCLPWNSAT---VLOQTYH--AHSRALQLGLGKHNYCRNPDN 107
D: 370 HGERYHGHVSKTRKIGICQWDATPHVPQISPTTPEAHL-----EKNYCENPDN 420
Q 108 -RRPWCYVQGLKPLVQECMVHDCADGKKPS 138
D: 421 DSHGFWCWTMDPTPTF-DYCAIKPCSGSAVPS 451

LT 48
Q 108 Q91XG8 PRELIMINARY; PRT; 716 AA.
A: Q91XG8;
D: 01-DEC-2001 (TremBLrel. 19, Created)
D: 01-DEC-2001 (TremBLrel. 19, Last sequence update)
D: 01-OCT-2001 (TremBLrel. 25, Last annotation update)
D: Hepatocyte growth factor-like.
D: MS11 OR HGFL.
O: Mus musculus (Mouse).
O: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
O: Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
O: NCBI_TaxID=10090;
E: [1]
E: SEQUENCE FROM N.A.
R: TISSUE=Liver;
R: Strausberg R.;
R: Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
C: -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
D: EMBL; BC010551; AAHL0551.1; -.
D: HSP; P00761; 1ANI.
D: MGD; MGI:96080; Met1.
D: GO; GO:0007566; P:embryo implantation; IC.
D: InterPro; IPR00903; Cys_Ser_trypsin.
D: InterPro; IPR000001; Kringle.
D: InterPro; IPR003014; PAN.
D: InterPro; IPR003609; Pan_app.
D: InterPro; IPR001254; Peptidase S1.
D: InterPro; IPR001314; Peptidase_S1A.
D: InterPro; IPR003966; Peptidase_S1A_pr.
D: Pfam; PF00051; kringle; 4.
D: Pfam; PF00024; PAN; 1.
D: Pfam; PF00089; trypsin; 1.
D: PRINTS; PR00722; CHYMOTRYPSIN.
D: PRINTS; PR00018; KRINGLE.
D: PRINTS; PR01505; PROTHROMBIN.
D: ProDom; PD000395; Kringle; 4.
D: SMART; SM00130; KR; 4.
D: SMART; SM00473; PAN_AP; 1.

DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS50070; KRINGLE_2; 4.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 716 AA; 80693 MW; 12474C48A7D4B46D CRC64;

Query Match 17.6%; Score 147.5; DB 11; Length 716;
Best Local Similarity 36.5%; Pred. No. 3.1e-07;
Matches 38; Conservative 11; Mismatches 34; Indels 21; Gaps 7;

QY 50 CYEGNGHFYRGKASDTMTGRPCLPWNSATVLQQTYYAHR-----SDALQLGLGKHNYCRN 104
DB 379 CYHSGEQYRGVSXKTRKGVQCOHWSSET-----PHKQFTTSAPOAGL-EANFCRN 430
QY 105 PD-NRRRPWCYVQVGLKP--LVQECMVHDCADGKKPS--SPPEE 143
DB 431 PDGDSHGWCYT---LDPEILFDYCALQRCDDDPSPSILDPDQ 471

RESULT 49
QYUIR6 PRELIMINARY; PRT; 113 AA.
ID QYUIR6;
AC QYUIR6;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Agolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21181705; PubMed-11285247;
RA Gogelkova W., Kraft H.G., Enholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158661; AAF03679.1; JOINED.
DR EMBL; AF158660; AAF03679.1; JOINED.
DR HSSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12697 MW; 51D4461D9C66312E CRC64;

Query Match 17.6%; Score 147; DB 4; Length 113;
Best Local Similarity 36.0%; Pred. No. 4.7e-08;
Matches 31; Conservative 14; Mismatches 33; Indels 8; Gaps 4;

QY 50 CYEGNGHFYRGKASDTMTGRPCLPWNSATVLQQTYYAHRSDALQLGLGKHNYCRNPDNR 109
DB 11 CYHGDGRSYRGISSTVTGTCTQSWSS--MIPHWHQRTPENYPNAGL-TENYCRNPDGSK 67
QY 110 RWCYVQVGLKPLV--QECMVHDCAD 133
DB 68 QPWCYT---TDPCVWEYCNLTQCSSE 90

RESULT 50
Q28398 PRELIMINARY; PRT; 2869 AA.
ID Q28398
```

AC Q28398;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Apolipoprotein(a) (Fragment).
OS Erinaceus europaeus (Western European hedgehog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceinae; Erinaceus.
OX NCBI_TaxID=9365;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=96025778; PubMed=7592597;
RA Lawn R.M., Boonmark N.W., Schwartz K., Lindahl G.E., Wade D.P.,
RA Byrne C.D., Fong K.J., Meer K., Patthy L.;
RT "The recurring evolution of lipoprotein(a). Insights from cloning of
RT hedgehog apolipoprotein(a).";
RL J. Biol. Chem. 270:24004-24009(1995).
CC -!- SIMILARITY: CONTAINS 31 KRINGLE DOMAINS.
DR EMBL; U33170; AAC48522.1; -.
DR PIR; T18518; T18518.
DR HSSP; P00747; 1PMK.
DR InterPro: IPR000001; Kringle.
DR Pfam: PF00051; kringle; 31.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 31.
DR SMART; SM00130; KR; 31.
DR PROSITE; PS00021; KRINGLE_1; 30.
DR PROSITE; PS00070; KRINGLE_2; 31.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER .1
SQ SEQUENCE 2869 AA; 318601 MW; 9527CEP985A4FE2A CRC64;
ary Match 17.6%; Score 147; DB 6; Length 2869;
st Local Similarity 29.3%; Pred. No. 1.6e-06;
ches 39; Conservative 11; Mismatches 61; Indels 22; Gaps 6;
QY 19 CYSNKYFSNIHNCNCPKFGGQHCEIDKSKT-----CYEGNGHFYRGKASTDTMGRP 70
Db 2558 CYTTSAMWEYCSIPA-----CESPTPEHLVVPQCLENGENYQGNVAITVSCQP 2611
QY 71 CLFWSATVLOQTYHAHRSDALQLGLGKHNCRPNRRPWCYVQGLKPLVQECMVHD 130
Db 2612 CQWRKQTPRHREYTPENYPKSL-PG--NYCRNPDGEIAPWCT-TNSAVRWEYCSIPT 2667
QY 131 CADGKFPSSPPE 143
Db 2668 C-----ESSSPPE 2676

Se: h completed: May 25, 2004, 14:57:18
Jo: ime : 24.6871 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

C protein - protein search, using sw model

R on: May 25, 2004, 14:43:00 ; Search time 34.9667 Seconds

(without alignments)
1155.508 Million cell updates/sec

T le: US-09-880-503-8

E fect score: 837

S uence: 1 SNEHQVPSNCDLNGGTCV.....QECNVHDCADKPKSPPEE 143

S ring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

S tched: 1586107 seqs, 282547505 residues

T al number of hits satisfying chosen parameters: 1586107

M imum DB seq length: 0

M imum DB seq length: 2000000000

E t-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

D atabase : A_Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

R	ult	No.	Score	Match	Length	ID	Description
1		837	100.0	143	5	AAR62991	Human uroki
2		837	100.0	337	4	AAR62998	Pro-uroki
3		837	100.0	337	5	AAR62993	Pro-uroki
4		837	100.0	411	1	AAR62926	Pro-uroki
5		837	100.0	411	2	AAR62936	Pro-uroki
6		837	100.0	411	2	AAR62936	Pro-uroki
7		837	100.0	411	2	AAR62936	Pro-uroki
8		837	100.0	411	2	AAR62936	Pro-uroki
9		837	100.0	411	2	AAR62936	Pro-uroki
10		837	100.0	411	2	AAR62936	Pro-uroki
11		837	100.0	411	2	AAR62936	Pro-uroki
12		837	100.0	411	2	AAR62936	Pro-uroki
13		837	100.0	411	2	AAR62936	Pro-uroki
14		837	100.0	411	2	AAR62936	Pro-uroki
15		837	100.0	411	2	AAR62936	Pro-uroki
16		837	100.0	411	2	AAR62936	Pro-uroki
17		837	100.0	411	2	AAR62936	Pro-uroki
18		837	100.0	411	2	AAR62936	Pro-uroki
19		837	100.0	411	2	AAR62936	Pro-uroki
20		837	100.0	411	2	AAR62936	Pro-uroki
21		837	100.0	411	2	AAR62936	Pro-uroki
22		837	100.0	411	2	AAR62936	Pro-uroki
23		837	100.0	411	2	AAR62936	Pro-uroki
24		837	100.0	411	2	AAR62936	Pro-uroki
25		837	100.0	411	2	AAR62936	Pro-uroki

ALIGNMENTS

RESULT 1

AAR62991 standard; protein; 143 AA.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

OS Homo sapiens.
 PN WO200197752-A2.
 XX 27-DEC-2001.
 XX 13-JUN-2001; 2001WO-US018976.
 XX 20-JUN-2000; 2000US-0212874P.
 XX (TYPE-) UNIV PENNSYLVANIA.
 XX Cines DB, Higazi AA;
 XX WPI; 2002-122240/16.
 XX N-PSDB; AAD27082.
 XX Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator.
 XX
 XX Claim 24; Fig.1H; 117pp; English.
 XX The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) amino
 CC terminal fragment (ATP) and connecting peptide
 XX
 XX Sequence 143 AA;
 XX
 XX Query Match 100.0%; Score 837; DB 5; Length 143;
 XX Best Local Similarity 100.0%; Pred. No. 9.9e-56;
 XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFNSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
 DB 1 SNEHQVPSNCDCLNGGTCVSNKYFNSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
 QY 61 KASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
 DB 61 KASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
 QY 121 PLVQECWVHDCADGKPKSPPEE 143
 DB 121 PLVQECWVHDCADGKPKSPPEE 143
 RE IT 2
 AA 5492
 ID AAG75492 standard; protein; 337 AA.
 XX AAG75492;
 XX
 XX 03-SEP-2001 (first entry)
 XX Human colon cancer antigen protein SEQ ID NO:6256.
 XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 10.
 XX Homo sapiens.
 XX

PN WO200122920-A2.
 XX 05-APR-2001.
 XX 28-SEP-2000; 2000WO-US026524.
 XX 29-SEP-1999; 99US-0157137P.
 XX 03-NOV-1999; 99US-0163280P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX WPI; 2001-235357/24.
 XX N-PSDB; AAH34897.
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers.
 XX
 XX Claim 11; Page 7707-7708; 9803pp; English.
 XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
 CC proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene therapy
 CC and vaccine production. N and P may be used in the prevention, diagnosis
 CC and treatment of diseases associated with inappropriate P expression. For
 CC example, N and P may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of P by expressing inactive proteins or to
 CC supplement the patient's own production of P. Additionally, N may be used
 CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the proteins. N and P
 CC can be used in the prevention, diagnosis and treatment of colorectal
 CC carcinomas and cancers. AAH37196 to AAH37204 and AAH77789 represent
 CC sequences used in the exemplification of the present invention. N.B.
 CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
 CC time of publication, meaning no sequences are present for SEQ ID NO.1027
 CC to 1032, 7921 and 7922
 XX
 XX Sequence 337 AA;
 XX
 XX Query Match 100.0%; Score 837; DB 4; Length 337;
 XX Best Local Similarity 100.0%; Pred. No. 2.1e-55;
 XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFNSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
 DB 27 SNEHQVPSNCDCLNGGTCVSNKYFNSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 86
 QY 61 KASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
 DB 87 KASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 146
 QY 121 PLVQECWVHDCADGKPKSPPEE 143
 DB 147 PLVQECWVHDCADGKPKSPPEE 169
 RESULT 3
 ABP41795
 ID ABP41795 standard; protein; 337 AA.
 XX
 XX AC ABP41795;
 XX
 XX 22-AUG-2002 (first entry)
 XX Human ovarian antigen HVUCB79, SEQ ID NO:2927.
 XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 XX

1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDSKTKCYEGNGHFYRG 60
61 KASTDTMGRPCLPWNATVLOQTYHAHRSDALQGLGKKNYCRNPNRRPWCYVQVGLK 120
61 KASTDTMGRPCLPWNATVLOQTYHAHRSDALQGLGKKNYCRNPNRRPWCYVQVGLK 120
121 PLVQECMVHDCADGKPKSPPEE 143
121 PLVQECMVHDCADGKPKSPPEE 143

UT 5
AA 5244
AAAR06244 standard; protein; 411 AA.
AC AAR06244;
DT 07-DEC-1990 (first entry)
XX Urokinase precursor protein.
XX Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis;
XX myocardial infarction.
XX Homo sapiens.
XX EP380334-A.
XX 01-AUG-1990.
XX 25-JAN-1990; 90EP-00300772.
XX 27-JAN-1989; 89JP-00016406.
XX 17-MAY-1989; 89JP-00121405.
XX (GREG) GREEN CROSS CORP.
XX Matsuda H, Ueda Y, Tamanouchi K;
XX WPI; 1990-233117/31.
XX Urokinase precursor-lipid composite - used as thrombolytic agent, having
XX prolonged half-life in the blood, enhanced bio-availability and improved
XX activity.
XX Claim 3; Fig 1; 11pp; English.
XX By forming a precursor-lipid composite, the half-life of this
XX thrombolytic agent in the blood may be increased, exhibiting improved
XX activity without abnormal acceleration of fibrinolytic activity. Compound
XX is useful as a thrombolytic agent in treatment of cerebral thrombosis,
XX myocardial infarction etc
XX Sequence 411 AA;

Query Match 100.0%; Score 837; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.4e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDSKTKCYEGNGHFYRG 60
1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDSKTKCYEGNGHFYRG 60
61 KASTDTMGRPCLPWNATVLOQTYHAHRSDALQGLGKKNYCRNPNRRPWCYVQVGLK 120
61 KASTDTMGRPCLPWNATVLOQTYHAHRSDALQGLGKKNYCRNPNRRPWCYVQVGLK 120
121 PLVQECMVHDCADGKPKSPPEE 143
121 PLVQECMVHDCADGKPKSPPEE 143

RESULT 6
AAR05117
ID AAR05117 standard; protein; 411 AA.
XX AC AAR05117;
XX 25-MAR-2003 (revised)
DT 04-OCT-1990 (first entry)
XX DE UK-S3 as encoded by PUKS3.
XX KW Urokinase; glycosylation.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Misc-difference 153 /label= synthetic mutation
FT /note= "old seq (Leu)"
FT Misc-difference 155 /label= synthetic mutation
FT /note= "old seq (Pro)"
XX PN EP370205-A.
XX PD 30-MAY-1990.
XX PF 28-SEP-1989; 89EP-00117981.
XX PR 29-SEP-1988; 88JP-00245705.
XX PA (KYOW) KYOWA HAKKO KOGYO KK.
XX PI Sasaki K, Nishi T, Yasumura S, Sato M, Itoh S;
XX WPI; 1990-165029/22.
XX N-PSDB; AAQ04486.
XX Polypeptide(s) with added carbohydrate chains - formed by modification of
XX aminoacid sequence, used to improve physio:chemical properties and/or
XX activities.
XX Disclosure; Page ?; 30pp; English.
XX The polypeptide is a deriv. of mature urokinase, designated UK-S3 which
XX has 2 amino acid substns. which result in an N-linked glycosylation site
XX giving the new protein improved stability and activity. See also AAR05113
XX -17; (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-
XX 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
XX Sequence 411 AA;
Query Match 100.0%; Score 837; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.4e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDSKTKCYEGNGHFYRG 60
1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDSKTKCYEGNGHFYRG 60
61 KASTDTMGRPCLPWNATVLOQTYHAHRSDALQGLGKKNYCRNPNRRPWCYVQVGLK 120
61 KASTDTMGRPCLPWNATVLOQTYHAHRSDALQGLGKKNYCRNPNRRPWCYVQVGLK 120
121 PLVQECMVHDCADGKPKSPPEE 143
121 PLVQECMVHDCADGKPKSPPEE 143

RESULT 7
AAW13634
ID AAW13634 standard; protein; 411 AA.
XX

cerebral thrombosis.
Homo sapiens.
EP405285-A.
02-JAN-1991.
18-JUN-1990; 90EP-00111471.
19-JUN-1989; 89JP-00156302.
(KYOW) KYOWA HAKKO KOGYO KK.
Yasamura S, Nishi T, Ito S;
WPI; 1991-008678/02.
N-PSDB; AAQ10170.
New plasminogen activator almost identical to natural pro-urokinase - is thrombin resistant and used for prophylaxis-treatment of cerebral thrombosis or myocardial infarction.
PS Disclosure; Page 9; 84pp; English.
XX UK-S3 is one example of a plasminogen activator which differs from natural human pro-urokinase at positions 153 and 155. (Leu substituted by Asn; Pro substituted by Thr, respectively). The derivative has decreased susceptibility to thrombin compared to natural type pro-UK and higher specific activity. See also AAQ10168 and AAQ10169
XX Sequence 411 AA;
Query Match 100.0%; Score 837; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.4e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGTGVSNKYFNSIHWNCNPKFQGOHCEIDKSKTCYEGNGHFYRG 60
DB 1 SNEHQVPSNCDCLNGTGVSNKYFNSIHWNCNPKFQGOHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRCPLPWNATVLQQTTHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRCPLPWNATVLQQTTHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKXSPPEE 143
DB 121 PLVQECMVHDCADGKXSPPEE 143
RE 121 PLVQECMVHDCADGKXSPPEE 143
AA 0225
ID AAR40225 standard; protein; 411 AA.
XX AAR40225;
AC AAR40225;
DT 10-FEB-1994 (first entry)
XX PUK.
XX Pre-urokinase; thrombolytic; blood; plasmid; PUK.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Region 10..60
FT /note= "Ser, Asn, Pro, Gly, or Tyr in the region 10 to 60 in the N-terminal of the human PUK can be replaced by Thr, Pro or Ala"
FT Domain 10..49
FT /label= EGF
FT Region 10..19
FT /label= Loop_1

Region 20..31
FT /label= Loop_2
FT Region 33..42
FT /label= Loop_3
FT Misc-difference 64
FT /note= "Thr encoded by AGT (sic)"
FT Misc-difference 177
FT /note= "Thr encoded by TAC (sic)"
XX JP05192142-A.
XX 03-AUG-1993.
XX 20-JAN-1992; 92JP-00030178.
XX 20-JAN-1992; 92JP-00030178.
XX (GREC) GREEN CROSS CORP.
XX WPI; 1993-277461/35.
XX N-PSDB; AAQ48228.
XX Mutant human pre-urokinase - by replacing specified aminoacid(s) in N-terminal for providing longer half-life in blood and higher thrombolytic ability.
XX Claim 1; Page 14-16; 26pp; Japanese.
XX Human pre-urokinase can be modified by replacing: Ser, Asn, Pro, Gly, or Tyr in the region ranging from residue 10 to 60 in the N-terminal of the human PUK by Thr, Pro or Ala. These mutants have a longer half-life in the blood and a higher thrombolytic ability. For examples see (AAR47956-R47960)
XX Sequence 411 AA;
Query Match 100.0%; Score 837; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.4e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGTGVSNKYFNSIHWNCNPKFQGOHCEIDKSKTCYEGNGHFYRG 60
DB 1 SNEHQVPSNCDCLNGTGVSNKYFNSIHWNCNPKFQGOHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRCPLPWNATVLQQTTHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRCPLPWNATVLQQTTHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKXSPPEE 143
DB 121 PLVQECMVHDCADGKXSPPEE 143
RESULT 11
AAR62997
ID AAR62997 standard; protein; 411 AA.
XX AAR62997;
AC AAR62997;
DT 25-MAR-2003 (revised)
DT 21-SEP-1995 (first entry)
XX Pro-urokinase mutant Gly306.
XX Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Gly306;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
KW systemic bleeding.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31

Disulfide-bond 33. .42
 Disulfide-bond 50. .131
 Disulfide-bond 71. .113
 Disulfide-bond 102. .136
 Disulfide-bond 148. .279
 Disulfide-bond 189. .205
 Disulfide-bond 197. .268
 Disulfide-bond 293. .362
 Domain 297. .313
 /note="flexible loop"
 Disulfide-bond 325. .341
 Disulfide-bond 352. .380
 WO9501427-A1
 12-JAN-1995.
 28-JUN-1994; 94WO-US007278.
 02-JUL-1993; 93US-00087163.
 (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 Liu J, Gurewich V;
 WPI; 1995-060991/08.
 Pro-urokinase mutants - have thrombolytic activity but reduced
 fibrinogenolysis activity and non-specific plasminogen activation.
 Claim 13; Fig 1; 46pp; English.
 AAR62991 is the wild type pro-urokinase, from which the new mutants
 described in AAR62992-R63008 were derived. These mutants retain the
 thrombolytic activity of the wild type protein, useful for the treatment
 of thromboembolism, but have a reduced fibrinogenolysis activity and non-
 specific plasminogen activation. The mutants can therefore be used for
 the lysis of fibrin clots without inducing systemic bleeding, as can be
 the case with the wild type protein. (Updated on 25-MAR-2003 to correct
 PN field.)
 Sequence 411 AA;
 Query Match 100.0%; Score 837; DB 2; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.4e-55;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFGQHCHEIDKSKTCYEGNGHFRG 60
 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFGQHCHEIDKSKTCYEGNGHFRG 60
 61 KASTDTMGPRCLPWSNATVLOQTYHAHRSALQGLGKHNCRPNDRRPPWCYVQVGLK 120
 61 KASTDTMGPRCLPWSNATVLOQTYHAHRSALQGLGKHNCRPNDRRPPWCYVQVGLK 120
 121 PLVQECMVHDCADGKPPSPPEE 143
 121 PLVQECMVHDCADGKPPSPPEE 143
 ULT 12
 63008
 AAR63008 standard; protein; 411 AA.
 AAR63008;
 25-MAR-2003 (revised)
 21-SEP-1995 (first entry)
 Pro-urokinase mutant Ser175 His187 His300 Ala301 His313.
 Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;
 non-specific plasminogen activation; systemic bleeding;

mutant Ser175 His187 His300 Ala301 His313.
 Homo sapiens.
 Key Location/Qualifiers
 Disulfide-bond 11. .19
 Disulfide-bond 13. .31
 Disulfide-bond 33. .42
 Disulfide-bond 50. .131
 Disulfide-bond 71. .113
 Disulfide-bond 102. .126
 Disulfide-bond 148. .279
 Disulfide-bond 189. .205
 Disulfide-bond 197. .268
 Disulfide-bond 293. .362
 Domain 297. .313
 /note="flexible loop"
 Disulfide-bond 325. .341
 Disulfide-bond 352. .380
 WO9501427-A1.
 12-JAN-1995.
 28-JUN-1994; 94WO-US007278.
 02-JUL-1993; 93US-00087163.
 (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 Liu J, Gurewich V;
 WPI; 1995-060991/08.
 Pro-urokinase mutants - have thrombolytic activity but reduced
 fibrinogenolysis activity and non-specific plasminogen activation.
 Claim 16; Fig 1; 46pp; English.
 AAR62991 is the wild type pro-urokinase, from which the new mutants
 described in AAR62992-R63008 were derived. These mutants retain the
 thrombolytic activity of the wild type protein, useful for the treatment
 of thromboembolism, but have a reduced fibrinogenolysis activity and non-
 specific plasminogen activation. The mutants can therefore be used for
 the lysis of fibrin clots without inducing systemic bleeding, as can be
 the case with the wild type protein. (Updated on 25-MAR-2003 to correct
 PN field.)
 Sequence 411 AA;
 Query Match 100.0%; Score 837; DB 2; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.4e-55;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFGQHCHEIDKSKTCYEGNGHFRG 60
 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFGQHCHEIDKSKTCYEGNGHFRG 60
 61 KASTDTMGPRCLPWSNATVLOQTYHAHRSALQGLGKHNCRPNDRRPPWCYVQVGLK 120
 61 KASTDTMGPRCLPWSNATVLOQTYHAHRSALQGLGKHNCRPNDRRPPWCYVQVGLK 120
 121 PLVQECMVHDCADGKPPSPPEE 143
 121 PLVQECMVHDCADGKPPSPPEE 143
 RESULT 13
 AAR62999
 ID AAR62999 standard; protein; 411 AA.
 XX
 AC AAR62999;
 XX

DT 25-MAR-2003 (revised)
 DT 21-SEP-1995 (first entry)
 XX Pro-urokinase mutant His313.
 DE
 XX Pro-urokinase; thrombolysis; fibrin clot lysis; mutant His313;
 KW reduced fibrinogenolysis; non-specific plasminogen activation;
 KW systemic bleeding.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FE Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Domain 297..313
 FT /note= "flexible loop"
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 FN WO9501427-A1.
 PN
 XX 12-JAN-1995.
 PD
 XX 28-JUN-1994; 94WO-US007278.
 PF
 XX 02-JUL-1993; 93US-00087163.
 XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Liu J, Gurewich V;
 DR WPI; 1995-060991/08.
 XX Pro-urokinase mutants - have thrombolytic activity but reduced
 PT fibrinogenolysis activity and non-specific plasminogen activation.
 XX Claim 11; Fig 1; 46pp; English.
 CC AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the treatment
 CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
 CC specific plasminogen activation. The mutants can therefore be used for
 CC the lysis of fibrin clots without inducing systemic bleeding, as can be
 CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
 CC PN field.)
 XX Sequence 411 AA;
 SQ
 Query Match 100.0%; Score 837; DB 2; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.4e-55;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIEDKSKTCYEGNGHFRG 60
 DI 1 SNEHQVPSNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIEDKSKTCYEGNGHFRG 60
 QY 61 KASTDTMGRPCLPWSATVLQOYTHAHRSDALQLGLGKHCYCRNPDRRPPWCYVQVGLK 120
 DI 61 KASTDTMGRPCLPWSATVLQOYTHAHRSDALQLGLGKHCYCRNPDRRPPWCYVQVGLK 120
 QY 121 PLVQECWHDCAKGPSPPEE 143
 DI 121 PLVQECWHDCAKGPSPPEE 143

RESULT 14
 AAR62994
 ID AAR62994 standard; protein; 411 AA.
 XX
 XX AAR62994;
 DT 25-MAR-2003 (revised)
 DT 21-SEP-1995 (first entry)
 XX
 DE Pro-urokinase mutant His299 Ala300.
 XX
 KW Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;
 KW non-specific plasminogen activation; systemic bleeding;
 KW mutant His 299 Ala300.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FE Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Domain 297..313
 FT /note= "flexible loop"
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 FN WO9501427-A1.
 PN
 XX 12-JAN-1995.
 PD
 XX 28-JUN-1994; 94WO-US007278.
 PF
 XX 02-JUL-1993; 93US-00087163.
 XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Liu J, Gurewich V;
 DR WPI; 1995-060991/08.
 XX Pro-urokinase mutants - have thrombolytic activity but reduced
 PT fibrinogenolysis activity and non-specific plasminogen activation.
 XX Claim 7; Fig 1; 46pp; English.
 CC AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the treatment
 CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
 CC specific plasminogen activation. The mutants can therefore be used for
 CC the lysis of fibrin clots without inducing systemic bleeding, as can be
 CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
 CC PN field.)
 XX Sequence 411 AA;
 SQ
 Query Match 100.0%; Score 837; DB 2; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.4e-55;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIEDKSKTCYEGNGHFRG 60
 DI 1 SNEHQVPSNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIEDKSKTCYEGNGHFRG 60
 QY 61 KASTDTMGRPCLPWSATVLQOYTHAHRSDALQLGLGKHCYCRNPDRRPPWCYVQVGLK 120
 DI 61 KASTDTMGRPCLPWSATVLQOYTHAHRSDALQLGLGKHCYCRNPDRRPPWCYVQVGLK 120
 QY 121 PLVQECWHDCAKGPSPPEE 143
 DI 121 PLVQECWHDCAKGPSPPEE 143

D 61 KASTDTMGRCPLPWSATVLTQTYHAHRSALQGLGKHNFCRNPDNRPRPWCYVQGLK 120
C 121 PLVQECMVHDCADGKKPSPPEE 143
D 121 PLVQECMVHDCADGKKPSPPEE 143
F ULT 15
A 53003
I AAR63003 standard; protein; 411 AA.
X AAR63003;
X 25-MAR-2003 (revised)
X 21-SEP-1995 (first entry)
X Pro-urokinase mutant Ser175 His187 His313.
X Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;
K non-specific plasminogen activation; systemic bleeding;
K mutant Ser175 His187 His313.
X Homo sapiens.
C
X
F Key Location/Qualifiers
F Disulfide-bond 11..19
F Disulfide-bond 13..31
F Disulfide-bond 33..42
F Disulfide-bond 50..131
F Disulfide-bond 71..113
F Disulfide-bond 102..126
F Disulfide-bond 148..279
F Disulfide-bond 189..205
F Disulfide-bond 197..268
F Disulfide-bond 283..362
F Domain 297..313
F Disulfide-bond 325..341
F Disulfide-bond 352..380
F WO9501427-A1
X 12-JAN-1995.
X 28-JUN-1994; 94WO-US007278.
X 02-JUL-1993; 93US-00087163.
X (NEWB-) NEW ENGLAND DEACONESS HOSPITAL.
X Liu J, Gurewich V;
X WPI; 1995-060991/08.
X Pro-urokinase mutants - have thrombolytic activity but reduced
F fibrinogenolysis activity and non-specific plasminogen activation.
X Claim 15; Fig 1; 46pp; English.
X AAR62991 is the wild type pro-urokinase, from which the new mutants
C described in AAR62992-R63008 were derived. These mutants retain the
C thrombolytic activity of the wild type protein, useful for the treatment
C of thromboembolism, but have a reduced fibrinogenolysis activity and non-
C specific plasminogen activation. The mutants can therefore be used for
C the lysis of fibrin clots without inducing systemic bleeding, as can be
C the case with the wild type protein. (Updated on 25-MAR-2003 to correct
C PN field.)
X Sequence 411 AA;
S Query Match 100.0%; Score 837; DB 2; Length 411;
Est Local Similarity 100.0%; Pred. No. 2.4e-55;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCEIDKSKTCYEGNGHFYRG 60
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCEIDKSKTCYEGNGHFYRG 60
Qy 61 KASTDTMGRCPLPWSATVLTQTYHAHRSALQGLGKHNFCRNPDNRPRPWCYVQGLK 120
Db 61 KASTDTMGRCPLPWSATVLTQTYHAHRSALQGLGKHNFCRNPDNRPRPWCYVQGLK 120
Qy 121 PLVQECMVHDCADGKKPSPPEE 143
Db 121 PLVQECMVHDCADGKKPSPPEE 143
RESULT 16
AAR63001
ID AAR63001 standard; protein; 411 AA.
XX AAR63001;
AC AAR63001;
XX 25-MAR-2003 (revised)
DT 21-SEP-1995 (first entry)
DT XX
DE Pro-urokinase mutant Ser175 His187 Ala313.
XX Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;
KW non-specific plasminogen activation; systemic bleeding;
KW mutant Ser175 His187 Ala313.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Domain 297..313
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
PN 12-JAN-1995.
XX PD
XX PF 28-JUN-1994; 94WO-US007278.
XX PR 02-JUL-1993; 93US-00087163.
XX PA (NEWB-) NEW ENGLAND DEACONESS HOSPITAL.
XX PI Liu J, Gurewich V;
XX WPI; 1995-060991/08.
XX Pro-urokinase mutants - have thrombolytic activity but reduced
FT fibrinogenolysis activity and non-specific plasminogen activation.
XX Claim 15; Fig 1; 46pp; English.
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the treatment
CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
CC specific plasminogen activation. The mutants can therefore be used for
CC the lysis of fibrin clots without inducing systemic bleeding, as can be
CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
CC the lysis of fibrin clots without inducing systemic bleeding, as can be

CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
 CC PN field.)
 XX Sequence 411 AA;
 SQ Query Match 100.0%; Score 837; DB 2; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.4e-55;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQPNSCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHGYRG 60
 DB 1 SNELHQPNSCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHGYRG 60
 QY 61 KASTDTMGRPCLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLK 120
 DB 61 KASTDTMGRPCLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLK 120
 QY 121 PLVQECMVHDCADGKPPSPPEE 143
 DB 121 PLVQECMVHDCADGKPPSPPEE 143
 RE UT 17
 AA 3006
 ID AAR63006 standard; protein; 411 AA.
 AC AAR63006;
 XX
 XX 25-MAR-2003 (revised)
 DT 21-SEP-1995 (first entry)
 DE Pro-urokinase mutant Ser175 His187 His300 Ala301 Ala313.
 XX
 XX Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;
 KW non-specific plasminogen activation; systemic bleeding;
 KV mutant Ser175 His187 His300 Ala301 Ala313.
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Domain 297..313
 FT /note= "flexible loop"
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 XX WO9501427-A1.
 XX 12-JAN-1995.
 XX 28-JUN-1994; 94WO-US007278.
 XX 02-JUL-1993; 93US-00087163.
 XX (NEWSE-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Liu J, Gurewicz V;
 XX WPI; 1995-060991/08.
 XX Pro-urokinase mutants - have thrombolytic activity but reduced
 XX fibrinogenolysis activity and non-specific plasminogen activation.
 XX Claim 16; Fig 1; 46pp; English.

XX AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the treatment
 CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
 CC specific plasminogen activation. The mutants can therefore be used for
 CC the lysis of fibrin clots without inducing systemic bleeding, as can be
 CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
 CC PN field.)
 XX Sequence 411 AA;
 SQ Query Match 100.0%; Score 837; DB 2; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.4e-55;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQPNSCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHGYRG 60
 DB 1 SNELHQPNSCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHGYRG 60
 QY 61 KASTDTMGRPCLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLK 120
 DB 61 KASTDTMGRPCLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLK 120
 QY 121 PLVQECMVHDCADGKPPSPPEE 143
 DB 121 PLVQECMVHDCADGKPPSPPEE 143
 RESULT 18
 AAR62992
 ID AAR62992 standard; protein; 411 AA.
 XX
 AC AAR62992;
 XX
 XX 25-MAR-2003 (revised)
 DT 21-SEP-1995 (first entry)
 DE Pro-urokinase mutant Ala300.
 XX
 XX Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala300;
 KW reduced fibrinogenolysis; non-specific plasminogen activation;
 KV systemic bleeding.
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Domain 297..313
 FT /note= "flexible loop"
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 XX WO9501427-A1.
 XX 12-JAN-1995.
 XX 28-JUN-1994; 94WO-US007278.
 XX 02-JUL-1993; 93US-00087163.
 XX (NEWSE-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Liu J, Gurewicz V;
 XX WPI; 1995-060991/08.
 XX Pro-urokinase mutants - have thrombolytic activity but reduced
 XX fibrinogenolysis activity and non-specific plasminogen activation.
 XX Claim 16; Fig 1; 46pp; English.

X: WPI; 1995-060991/08.
D: Pro-urokinase mutants - have thrombolytic activity but reduced
F: fibrinogenolysis activity and non-specific plasminogen activation.
P: Claim 5; Fig 1; 46pp; English.
X: AAR62991 is the wild type pro-urokinase, from which the new mutants
C: described in AAR62992-R63008 were derived. These mutants retain the
C: thrombolytic activity of the wild type protein, useful for the treatment
C: of thromboembolism, but have a reduced fibrinogenolysis activity and non-
C: specific plasminogen activation. The mutants can therefore be used for
C: the lysis of fibrin clots without inducing systemic bleeding, as can be
C: the case with the wild type protein. (Updated on 25-MAR-2003 to correct
C: PN field.)
X: Sequence 411 AA;
S: Query Match 100.0%; Score 837; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.4e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Q: 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
D: 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
Q: 61 KASDTMGPRCLPWSNATVLOQTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVOVGLK 120
D: 61 KASDTMGPRCLPWSNATVLOQTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVOVGLK 120
Q: 121 PLVQECMVHDCADGKKPSSPPEE 143
D: 121 PLVQECMVHDCADGKKPSSPPEE 143
R: LT 19
A: 3002
I: AAR63002 standard; protein; 411 AA.
X: AAR63002;
X: 25-MAR-2003 (revised)
D: 21-SEP-1995 (first entry)
X: Pro-urokinase mutant Ser175 His187 Gly306.
X: Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;
X: non-specific plasminogen activation; systemic bleeding;
X: mutant Ser175 His187 Gly306.
O: Homo sapiens.
X: Key Location/Qualifiers
F: Disulfide-bond 11..19
F: Disulfide-bond 13..31
F: Disulfide-bond 33..42
F: Disulfide-bond 50..131
F: Disulfide-bond 71..113
F: Disulfide-bond 102..126
F: Disulfide-bond 148..279
F: Disulfide-bond 189..205
F: Disulfide-bond 197..268
F: Disulfide-bond 293..362
F: Domain 297..313
F: Disulfide-bond 325..341
F: Disulfide-bond 352..380
X: WO9501427-A1
X: 12-JAN-1995.

PF 28-JUN-1994; 94WO-US007278.
XX
PR 02-JUL-1993; 93US-00087163.
PA (NEWB-) NEW ENGLAND DEACONESS HOSPITAL.
XX
PI Liu J, Gurewich V;
XX
DR WPI; 1995-060991/08.
XX
PT Pro-urokinase mutants - have thrombolytic activity but reduced
PT fibrinogenolysis activity and non-specific plasminogen activation.
XX
PS Claim 15; Fig 1; 46pp; English.
XX
CC AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the treatment
CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
CC specific plasminogen activation. The mutants can therefore be used for
CC the lysis of fibrin clots without inducing systemic bleeding, as can be
CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
CC PN field.)
XX
SQ Sequence 411 AA;
Query Match 100.0%; Score 837; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.4e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Q: 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
D: 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
Q: 61 KASDTMGPRCLPWSNATVLOQTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVOVGLK 120
D: 61 KASDTMGPRCLPWSNATVLOQTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVOVGLK 120
Q: 121 PLVQECMVHDCADGKKPSSPPEE 143
D: 121 PLVQECMVHDCADGKKPSSPPEE 143
RESULT 20
AAR63007
ID AAR63007 standard; protein; 411 AA.
XX
AC AAR63007;
XX
DT 25-MAR-2003 (revised)
DT 21-SEP-1995 (first entry)
XX
DE Pro-urokinase mutant Ser175 His187 Ala300 Ala301 His313.
XX
KW Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;
KW non-specific plasminogen activation; systemic bleeding;
KW mutant Ser175 His187 Ala300 Ala301 His313.
XX
OS Homo sapiens.
X: Key Location/Qualifiers
F: Disulfide-bond 11..19
F: Disulfide-bond 13..31
F: Disulfide-bond 33..42
F: Disulfide-bond 50..131
F: Disulfide-bond 71..113
F: Disulfide-bond 102..126
F: Disulfide-bond 148..279
F: Disulfide-bond 189..205
F: Disulfide-bond 197..268
F: Disulfide-bond 293..362
F: Domain 297..313
F: /note= "flexible loop"

FT Disulfide-bond 325. .341
 FT Disulfide-bond 352. .380
 XX WO9501427-A1.
 XX 12-JAN-1995.
 XX 28-JUN-1994; 94WO-US007278.
 XX 02-JUL-1993; 93US-00087163.
 XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Liu J, Gurewich V;
 XX WPI; 1995-060991/08.
 XX Pro-urokinase mutants - have thrombolytic activity but reduced
 FT fibrinogenolysis activity and non-specific plasminogen activation.
 XX Claim 16; Fig 1; 46pp; English.
 XX AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the treatment
 CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
 CC specific plasminogen activation. The mutants can therefore be used for
 CC the lysis of fibrin clots without inducing systemic bleeding, as can be
 CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
 CC PN field.)
 XX Sequence 411 AA;
 XX Query Match 100.0%; Score 837; DB 2; Length 411;
 XX Best Local Similarity 100.0%; Pred. No. 2.4e-55;
 XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNEHQPVSNCDCCLNGTGVSNKYFNSNIHWNCNPKKFGGQHCIDKSKTCYEGNGHYRG 60
 DB 1 SNEHQVPSNCDCLNGTGVSNKYFNSNIHWNCNPKKFGGQHCIDKSKTCYEGNGHYRG 60
 QY 61 KASTDTMGRCPLPWSATVLQQTTHAHRSDALQLGLGKHNCRPNPNRRPWCYVQVGLK 120
 DB 61 KASTDTMGRCPLPWSATVLQQTTHAHRSDALQLGLGKHNCRPNPNRRPWCYVQVGLK 120
 QY 121 PLVQECNVHDCADGKPPSPPEE 143
 DB 121 PLVQECNVHDCADGKPPSPPEE 143
 RE DT 21
 AA 2995
 ID AAR62995 standard; protein; 411 AA.
 XX AAR62995;
 XX 25-MAR-2003 (revised)
 XX 21-SEP-1995 (first entry)
 XX Pro-urokinase mutant Ala300 Ala301.
 XX Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;
 KW non-specific plasminogen activation; mutant Ala300 Ala301;
 KW systemic bleeding.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT Disulfide-bond 11. .19
 FT Disulfide-bond 13. .31
 FT Disulfide-bond 33. .42
 FT Disulfide-bond 50. .131
 FT Disulfide-bond 71. .113

FT Disulfide-bond 102. .126
 FT Disulfide-bond 148. .279
 FT Disulfide-bond 189. .205
 FT Disulfide-bond 197. .268
 FT Disulfide-bond 293. .362
 FT Domain 297. .313
 FT /note= "flexible loop"
 FT Disulfide-bond 325. .341
 FT Disulfide-bond 352. .380
 XX WO9501427-A1.
 XX 12-JAN-1995.
 XX 28-JUN-1994; 94WO-US007278.
 XX 02-JUL-1993; 93US-00087163.
 XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Liu J, Gurewich V;
 XX WPI; 1995-060991/08.
 XX Pro-urokinase mutants - have thrombolytic activity but reduced
 FT fibrinogenolysis activity and non-specific plasminogen activation.
 XX Claim 9; Fig 1; 46pp; English.
 XX AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the treatment
 CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
 CC specific plasminogen activation. The mutants can therefore be used for
 CC the lysis of fibrin clots without inducing systemic bleeding, as can be
 CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
 CC PN field.)
 XX Sequence 411 AA;
 XX Query Match 100.0%; Score 837; DB 2; Length 411;
 XX Best Local Similarity 100.0%; Pred. No. 2.4e-55;
 XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNEHQPVSNCDCCLNGTGVSNKYFNSNIHWNCNPKKFGGQHCIDKSKTCYEGNGHYRG 60
 DB 1 SNEHQVPSNCDCLNGTGVSNKYFNSNIHWNCNPKKFGGQHCIDKSKTCYEGNGHYRG 60
 QY 61 KASTDTMGRCPLPWSATVLQQTTHAHRSDALQLGLGKHNCRPNPNRRPWCYVQVGLK 120
 DB 61 KASTDTMGRCPLPWSATVLQQTTHAHRSDALQLGLGKHNCRPNPNRRPWCYVQVGLK 120
 QY 121 PLVQECNVHDCADGKPPSPPEE 143
 DB 121 PLVQECNVHDCADGKPPSPPEE 143
 RESULT 22
 AAR63000
 ID AAR63000 standard; protein; 411 AA.
 XX AAR63000;
 XX 25-MAR-2003 (revised)
 XX 21-SEP-1995 (first entry)
 XX Pro-urokinase mutant Ser175 His187.
 XX Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;
 KW non-specific plasminogen activation; systemic bleeding;
 KW mutant Ser175 His187.
 XX Homo sapiens.
 OS

ID XX AAR62996 standard; protein; 411 AA.
 AC AAR62996;
 XX 25-MAR-2003 (revised)
 DT 21-SEP-1995 (first entry)
 XX Pro-urokinase mutant His300 Ala301.
 DE Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;
 XX non-specific plasminogen activation; mutant His300 Ala301;
 KW systemic bleeding.
 XX Homo sapiens.
 OS
 XX
 FE Key Location/Qualifiers
 FT Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..205
 FT Disulfide-bond 189..268
 FT Disulfide-bond 293..362
 FT Domain 297..313
 FT /note= "flexible loop"
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 XX WO9501427-A1.
 XX 12-JAN-1995.
 XX 28-JUN-1994; 94WO-US007278.
 XX 02-JUL-1993; 93US-00087163.
 XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Liu J, Gurewich V;
 XX WPI; 1995-060991/08.
 XX Pro-urokinase mutants - have thrombolytic activity but reduced
 FT fibrinogenolysis activity and non-specific plasminogen activation.
 XX Claim 9; Fig 1; 46pp; English.
 CC AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the treatment
 CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
 CC specific plasminogen activation. The mutants can therefore be used for
 CC the lysis of fibrin clots without inducing systemic bleeding, as can be
 CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
 CC PN field.)
 XX Sequence 411 AA;
 XX
 XX Query Match 100.0%; Score 837; DB 2; Length 411;
 XX Best Local Similarity 100.0%; Pred. No. 2.4e-55;
 XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEIDKSKTCYEGNGHFRG 60
 DL 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEIDKSKTCYEGNGHFRG 60
 QY 61 KASTDTMGPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 DL 61 KASTDTMGPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120

QY 121 PLVQECMVHDCADGKXPSPPEE 143
 DB 121 PLVQECMVHDCADGKXPSPPEE 143
 RESULT 25
 AAR63005
 ID AAR63005 standard; protein; 411 AA.
 XX AAR63005;
 AC AAR63005;
 XX 25-MAR-2003 (revised)
 DT 21-SEP-1995 (first entry)
 XX Pro-urokinase mutant Ser175 His187 Ala300 Ala301 Ala313.
 DE Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;
 KW non-specific plasminogen activation; systemic bleeding;
 XX mutant Ser175 His187 Ala300 Ala301 Ala313.
 OS Homo sapiens.
 XX
 FE Key Location/Qualifiers
 FT Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..205
 FT Disulfide-bond 189..268
 FT Disulfide-bond 293..362
 FT Domain 297..313
 FT /note= "flexible loop"
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 XX WO9501427-A1.
 XX 12-JAN-1995.
 XX 28-JUN-1994; 94WO-US007278.
 XX 02-JUL-1993; 93US-00087163.
 XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Liu J, Gurewich V;
 XX WPI; 1995-060991/08.
 XX Pro-urokinase mutants - have thrombolytic activity but reduced
 FT fibrinogenolysis activity and non-specific plasminogen activation.
 XX Claim 16; Fig 1; 46pp; English.
 CC AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the treatment
 CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
 CC specific plasminogen activation. The mutants can therefore be used for
 CC the lysis of fibrin clots without inducing systemic bleeding, as can be
 CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
 CC PN field.)
 XX Sequence 411 AA;
 XX
 XX Query Match 100.0%; Score 837; DB 2; Length 411;
 XX Best Local Similarity 100.0%; Pred. No. 2.4e-55;
 XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEIDKSKTCYEGNGHFRG 60

1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFRG 60
61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHCYCRNPDRRRPWCYVQVGLK 120
61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHCYCRNPDRRRPWCYVQVGLK 120
121 PLVQECMVHDCADGKKPSSPPEE 143
121 PLVQECMVHDCADGKKPSSPPEE 143

LT 26
52991
AAR62991 standard; protein; 411 AA.
AAR62991;
25-MAR-2003 (revised)
21-SEP-1995 (first entry)
Pro-urokinase.
Pro-urokinase; thrombolysis; fibrin clot lysis.
Homo sapiens.
Key Location/Qualifiers
Disulfide-bond 11..19
Disulfide-bond 13..31
Disulfide-bond 33..42
Disulfide-bond 50..131
Disulfide-bond 71..113
Disulfide-bond 102..126
Disulfide-bond 148..279
Disulfide-bond 189..205
Disulfide-bond 197..268
Disulfide-bond 293..362
Domain 297..313
Disulfide-bond 325..341
Disulfide-bond 352..380
WO9501427-A1.
12-JAN-1995.
28-JUN-1994; 94WO-US007278.
02-JUL-1993; 93US-00087163.
(NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
Liu J, Gurewich V;
WPI; 1995-060991/08.
Pro-urokinase mutants - have thrombolytic activity but reduced fibrinogenolysis activity and non-specific plasminogen activation.
Disclosure; Fig 1; 46pp; English.
AAR62991 is the wild type pro-urokinase, from which the new mutants described in AAR62992-R63008 were derived. These mutants retain the thrombolytic activity of the wild type protein, useful for the treatment of thromboembolism, but have a reduced fibrinogenolysis activity and non-specific plasminogen activation. The mutants can therefore be used for the lysis of fibrin clots without inducing systemic bleeding, as can be the case with the wild type protein. (Updated on 25-MAR-2003 to correct PN field.)
Sequence 411 AA;

Query Match 100.0%; Score 837; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.4e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFRG 60
Db 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFRG 60
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHCYCRNPDRRRPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHCYCRNPDRRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 121 PLVQECMVHDCADGKKPSSPPEE 143

RESULT 27
AAR62998
ID AAR62998 standard; protein; 411 AA.
XX AAR62998;
XX 25-MAR-2003 (revised)
DT 21-SEP-1995 (first entry)
XX Pro-urokinase mutant Ala313.
XX Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala313;
XX reduced fibrinogenolysis; non-specific plasminogen activation;
XX systemic bleeding.
XX Homo sapiens.
XX Key Location/Qualifiers
Disulfide-bond 11..19
Disulfide-bond 13..31
Disulfide-bond 33..42
Disulfide-bond 50..131
Disulfide-bond 71..113
Disulfide-bond 102..126
Disulfide-bond 148..279
Disulfide-bond 189..205
Disulfide-bond 197..268
Disulfide-bond 293..362
Domain 297..313
Disulfide-bond 325..341
Disulfide-bond 352..380
WO9501427-A1.
12-JAN-1995.
28-JUN-1994; 94WO-US007278.
02-JUL-1993; 93US-00087163.
XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX Liu J, Gurewich V;
XX WPI; 1995-060991/08.
XX Pro-urokinase mutants - have thrombolytic activity but reduced
PT fibrinogenolysis activity and non-specific plasminogen activation.
XX Claim 1; Fig 1; 46pp; English.
CC AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the treatment
CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-

specific plasminogen activation. The mutants can therefore be used for the lysis of fibrin clots without inducing systemic bleeding, as can be the case with the wild type protein. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 411 AA;

Query Match 100.0%; Score 837; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.4e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNEHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
Db 1 SNEHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
Qy 61 KASTDTMGRCLPWSATVLOQTYHAHRSDALQGLGKHNCRPNRRPWCYVQVGLK 120
Db 61 KASTDTMGRCLPWSATVLOQTYHAHRSDALQGLGKHNCRPNRRPWCYVQVGLK 120
Qy 121 PLVQECMVHDCADGKPPSPPEE 143
Db 121 PLVQECMVHDCADGKPPSPPEE 143

RE 28

AA 1993

ID AAR62993 standard; protein; 411 AA.

AC AAR62993;

XX 25-MAR-2003 (revised)

DT 21-SEP-1995 (first entry)

XX Pro-urokinase mutant His300.

DE XX

Pro-urokinase; thrombolysis; fibrin clot lysis; mutant His300;
reduced fibrinogenolysis; non-specific plasminogen activation;
systemic bleeding.

XX Homo sapiens.

Key Location/Qualifiers

Disulfide-bond 11..19
Disulfide-bond 13..31
Disulfide-bond 33..42
Disulfide-bond 50..131
Disulfide-bond 71..113
Disulfide-bond 102..126
Disulfide-bond 148..279
Disulfide-bond 189..205
Disulfide-bond 197..268
Disulfide-bond 293..362
Domain 297..313
/note= "flexible loop"

Disulfide-bond 325..341

Disulfide-bond 352..380

WO9501427-A1.

12-JAN-1995.

28-JUN-1994; 94WO-US007278.

02-JUL-1993; 93US-00087163.

(NEWB-) NEW ENGLAND DEACONESS HOSPITAL.
Liu J, Gurewich V;

WPI; 1995-060991/08.

Pro-urokinase mutants - have thrombolytic activity but reduced
fibrinogenolysis activity and non-specific plasminogen activation.

XX Claim 5; Fig 1; 46pp; English.

PS AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the treatment
CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
CC specific plasminogen activation. The mutants can therefore be used for
CC the lysis of fibrin clots without inducing systemic bleeding, as can be
CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
CC PN field.)

XX Sequence 411 AA;

Query Match 100.0%; Score 837; DB 2; Length 411;

Best Local Similarity 100.0%; Pred. No. 2.4e-55;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNEHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60

Db 1 SNEHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60

Qy 61 KASTDTMGRCLPWSATVLOQTYHAHRSDALQGLGKHNCRPNRRPWCYVQVGLK 120

Db 61 KASTDTMGRCLPWSATVLOQTYHAHRSDALQGLGKHNCRPNRRPWCYVQVGLK 120

Qy 121 PLVQECMVHDCADGKPPSPPEE 143

Db 121 PLVQECMVHDCADGKPPSPPEE 143

RESULT 29

AAR62926

ID AAR62926 standard; protein; 411 AA.

XX

AC AAR62926;

XX 03-AUG-1996 (first entry)

DT

DE Pro-urokinase.

XX Pro-urokinase; plasminogen activator; fusion drug; drug delivery;
platelet; cardiovascular disease; thrombolytic.

XX Homo sapiens.

Key Location/Qualifiers

Region 1..132
Domain 1..45
/label= A-chain
Disulfide-bond 11
/label= Growth_factor_domain
Disulfide-bond 13
/note= "disulfide between Cys11 and Cys19"
Disulfide-bond 33
/note= "disulfide bond between Cys13 and Cys31"
Disulfide-bond 46..132
/note= "disulfide bond between Cys33 and Cys42"
Domain 46..132
/label= Kringle_domain
Disulfide-bond 50
/note= "disulfide bond between Cys50 and Cys131"
Disulfide-bond 71
/note= "disulfide bond between Cys71 and Cys113"
Disulfide-bond 102
/note= "disulfide bond between Cys102 and Cys126"
Region 133..158
/label= Linker_region
Disulfide-bond 148
/note= "disulfide bond between Cys148 and Cys279"
Cleavage-site 156..157
/note= "thrombin cleavage site"
Cleavage-site 158..159
/note= "plasmin cleavage site"

F Region 159..411
F /label= B-chain
F Disulfide-bond 189
F /note= "disulfide bond between Cys189 and Cys205"
F Disulfide-bond 197
F /note= "disulfide bond between Cys197 and Cys268"
F Disulfide-bond 293
F /note= "disulfide bond between Cys293 and Cys362"
F Disulfide-bond 325
F /note= "disulfide bond between Cys325 and Cys341"
F Disulfide-bond 352
F /note= "disulfide bond between Cys352 and Cys380"
X WO9604004-A1.
X 15-FEB-1996.
X 03-AUG-1995; 95WO-US009848.
X 05-AUG-1994; 94US-00286748.
X (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
X Gurewich V;
X WPI; 1996-129123/13.
X N-PSDB; AAT18237.
P Fusion product of plasminogen activator A chain and drug - targeted to
P platelets, useful for treatment of cardiovascular disease.
P Claim 3; Page 39-40; 61pp; English.
C A portion of the A-chain (pref. amino acids 1-132) of pro-urokinase
C (AAR92926) can be linked to a drug, e.g. hirudin, somatostatin or their
C analogues (see also AAR92925 and AAR92927-33) and used in the prodn. of
C new fusion drugs. The constructs can be obt'd. by expression of the
C appropriate nucleotide sequences in transformed host cells. When
C administered to a patient, the A-chain binds the fusion drug to the
C platelet outer membrane, i.e. to the site of thrombosis or vascular
C injury. Cleavage sites for thrombin and/or plasmin with the fusion drug
C allow the release of the drug at the target site
S Sequence 411 AA;
Query Match 100.0%; Score 837; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.4e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
C 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEDKSKTCYEGNGHFYRG 60
D 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEDKSKTCYEGNGHFYRG 60
Q 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVLK 120
D 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVLK 120
C 121 PLVQECMVHDCADGKPKSSPPEE 143
D 121 PLVQECMVHDCADGKPKSSPPEE 143
F ULT 30
A 92836
I AAY92836 standard; protein; 411 AA.
X AAY92836;
X 29-AUG-2000 (first entry)
D Urokinase plasminogen activator (uPA).
X N-terminal; pro-uPA; urokinase plasminogen activator; receptor; uPAR;

KW anti-cancer; anti-metastatic; anti-proliferative; anti-atherosclerotic;
KW anti-thrombotic; anti-angiogenic; anti-inflammatory; anti-arthritis;
KW anti-fibrotic; apoptotic; vasotrophic; anti-diabetic; ophthalmological;
KW thrombolytic.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..203
FT Disulfide-bond 187..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO200026353-A1.
XX 11-MAY-2000.
XX 28-OCT-1999; 99WO-US025210.
XX 29-OCT-1998; 98US-00181816.
XX (ANGS-) ANGSTROM PHARM INC.
XX Mazar AP, Jones TR;
XX WPI; 2000-365605/31.
XX New cyclic peptide, useful for treatment or diagnosis of e.g. tumors and
XX other diseases involving cell proliferation or migration, targets the
XX urokinase plasminogen activator receptor.
XX Disclosure; Fig 1; 93pp; English.
SS The present sequence shows the wild-type urokinase plasminogen activator
CC (uPA). Cyclic peptides based on the amino acids residues 20-30 (the
CC receptor-binding region) of uPA are claimed. These cyclic peptides target
CC the uPA receptor (uPAR), allowing therapeutic or diagnostic agents to be
CC delivered to uPAR-expressing cells. The cyclic peptides are used,
CC optionally when linked to a therapeutic agent, to inhibit migration,
CC invasion and proliferation of cells, or angiogenesis, or to induce
CC apoptosis. Particularly they are used, in human or veterinary medicine,
CC to treat diseases characterized by these processes, e.g. solid tumors,
CC leukaemia or lymphoma (or their metastases); benign hyperplasia;
CC atherosclerosis; restenosis; ischaemia; deep vein thrombosis; neovascular
CC glaucoma; diabetic retinopathy; arthritis; fibrosis; bone fracture etc.,
CC most particularly growth, invasion and metastasis of tumors. When
CC labeled, the cyclic peptides can be used for diagnostic detection of uPAR
CC (a marker of metastasis) on cells, tissues etc., in vivo or in vitro, and
CC when immobilized they are used to isolate uPAR or cells that express
CC them. The cyclic peptides are stable, soluble in water, bind strongly to
CC uPAR, are relatively inexpensive to produce and may be derivatized by
CC attachment of therapeutic or diagnostic agents without significantly
CC affecting their binding. Since they target uPAR, they should have
CC relatively low systemic toxicity and only low doses are required
XX Sequence 411 AA;
Query Match 100.0%; Score 837; DB 3; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.4e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKKNYCRNPNRRRPPWCYVOVGLK 120
 DE 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKKNYCRNPNRRRPPWCYVOVGLK 120
 QY 121 PLVQECWVHDCADGKPKSPPEE 143
 DB 121 PLVQECWVHDCADGKPKSPPEE 143
 RE JT 31
 AA 544
 IL AAE16544 standard; protein; 411 AA.
 XX AAE16544;
 AC
 DT 09-APR-2002 (first entry)
 XX Human urokinase-type plasminogen activator tcupA and scuPA protein.
 DE Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
 KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
 KW clotting disorder; uterine contraction disorder; respiratory disease;
 KW male impotence; adult respiratory distress syndrome; tcupA; scuPA;
 KW two chain urokinase; single chain urokinase.
 XX Homo sapiens.
 OE WO200197752-A2.
 XX
 PN 27-DEC-2001.
 XX
 PD 13-JUN-2001; 2001WO-US018976.
 XX
 PF 20-JUN-2000; 2000US-0212874P.
 XX
 PF (UYPE-) UNIV PENNSYLVANIA.
 XX
 PA Cines DB, Higazi AA;
 XX
 PI WPI; 2002-122240/16.
 XX
 DE N-PSDB; AAD27077.
 XX
 DR Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator.
 XX
 PS Claim 9; Fig 1C; 117pp; English.
 XX
 CC The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic microangiopathies, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) two chain
 CC urokinase (tcupA) and single chain urokinase (scuPA) protein
 XX
 SQ Sequence 411 AA;
 Query Match 100.0%; Score 837; DB 5; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.4e-55;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQPNSCDCLNGGTCVSNKYFNIHWNCNPKFGGQHCEIDSKTCYEGNGHFYRG 60

DB 1 SNELHQPNSCDCLNGGTCVSNKYFNIHWNCNPKFGGQHCEIDSKTCYEGNGHFYRG 60
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKKNYCRNPNRRRPPWCYVOVGLK 120
 DB 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKKNYCRNPNRRRPPWCYVOVGLK 120
 QY 121 PLVQECWVHDCADGKPKSPPEE 143
 DB 121 PLVQECWVHDCADGKPKSPPEE 143
 RESULT 32
 AAR10334
 ID AAR10334 standard; protein; 412 AA.
 XX
 AC AAR10334;
 DT 25-MAR-2003 (revised)
 DT 05-APR-1991 (first entry)
 XX
 XX Recombinant single-chain urine plasminogen activator.
 DE Single chain urine plasminogen activator; enterobacteriaceae; scu-PA;
 KW cardiac infarction; pulmonary embolism; arterial occlusive disease.
 XX
 OS Synthetic.
 XX
 XX EP408945-A.
 PN
 XX 23-JAN-1991.
 PD
 XX 29-JUN-1990; 90EP-00112400.
 PF
 XX 19-JUL-1989; 89DE-03923866.
 PR
 XX (CHEP) GRUENENTHAL GMBH.
 PA
 XX Brigelius R, Flohe L, Hillen W, Steffens GJ, Straassburg W;
 PI Wilhelm M;
 PI
 XX WPI; 1991-023590/04.
 DR N-PSDB; AAQ10282.
 DR
 XX New plasmids for expressing urine plasminogen activator - in
 PT enterobacteriaceae have operon contg. promoter, ribosome binding site,
 PT synthetic gene and terminator, etc.
 XX
 PS Disclosure; Fig 15; 44pp; German.
 XX
 CC This recombinant single-chain urine-plasminogen activator (scu-PA), is
 CC encoded by a structural gene inserted into plasmid pBR322 (from which the
 CC nic/bom and/or tet resistance genes have been removed), following
 CC insertion of a multiple cloning site (see AAQ10281), a terminator and a
 CC synthetic Trp promoter. The resultant expression construct can be used to
 CC produce high yields of this scu-PA protein which can be refolded to
 CC therapeutically active recombinant scu-PA. This is used in the treatment
 CC of e.g. cardiac infarction, pulmonary embolism or arterial occlusive
 CC disease. (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 412 AA;
 Query Match 100.0%; Score 837; DB 2; Length 412;
 Best Local Similarity 100.0%; Pred. No. 2.5e-55;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQPNSCDCLNGGTCVSNKYFNIHWNCNPKFGGQHCEIDSKTCYEGNGHFYRG 60
 DB 2 SNELHQPNSCDCLNGGTCVSNKYFNIHWNCNPKFGGQHCEIDSKTCYEGNGHFYRG 61
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKKNYCRNPNRRRPPWCYVOVGLK 120
 DB 62 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKKNYCRNPNRRRPPWCYVOVGLK 121

C 121 PLVQECMVHDCADGKKPSSPPEE 143
D 122 PLVQECMVHDCADGKKPSSPPEE 144

E ULT 33
A 24579 AAW24579 standard; protein; 424 AA.
X AAW24579;
X 25-MAR-2003 (revised)
D 11-NOV-1997 (first entry)
X Inhibitor resistant modified urokinase.

X Urokinase; inhibitor resistant; plasminogen; human; whey acid protein;
K plasminogen activator inhibitor-1; plasminogen activator; protease; WAP;
K plasmin; proteolytic enzyme; clot lysis; heart attack; heart muscle.
X Homo sapiens.
C Synthetic.

X Key Location/Qualifiers
F Peptide 1..19
F /note= "whey acid protein signal peptide"
F Protein 20..430
F /note= "urokinase"
F Domain 20..64
F /note= "E-domain"
F Domain 61..150
F /note= "Kringle-1 domain"
F Domain 179..424
F /note= "P-domain"
F Misc-difference 197..198
F /note= "site of 6 residue (Arg-His-Arg-Gly-Ser)
F deletion"

X US5648253-A.
X 15-JUL-1997.
X 08-SEP-1992; 92US-00942157.
X 20-DEC-1990; 90US-00631673.
X (TSIT-) TSI CORP.
X Wei C;
X WPI; 1997-372062/34.
X N-PSDB; AAT80076.

X Deletion-modified urokinase protein - with increased resistance to
X inhibition by plasminogen activator inhibitor-1.
X Claim 2; Col; 16pp; English.

X This sequence represents the modified full length urokinase of the
C invention, including the whey acid protein (WAP) signal peptide. This
C sequence had residues 179-184 of the wild type urokinase sequence (see
C AAW24578) deleted. This modified urokinase cleaves plasminogen, and has a
C lower binding affinity for plasminogen activator inhibitor-1 than the
C corresponding unmodified urokinase. Urokinase is one of two types of
C mammalian plasminogen activators (PA), the other being tissue type PA.
C PAs catalyze the conversion of the circulating zymogen plasminogen to the
C broad spectrum protease plasmin by limited proteolysis. The modified
C urokinase can be used for clot lysis, specifically to dissolve heart
C attack-causing clots before they cause permanent damage to heart muscle.
C The urokinase mutant is more resistant to inhibition by plasminogen
C activator inhibitors than the unmodified urokinase. It can be selectively
C expressed and secreted from the mammary glands of transgenic animals.

CC (Updated on 25-MAR-2003 to correct PF field.)
XX Sequence 424 AA;
SQ

Query Match 100.0%; Score 837; DB 2; Length 424;
Best Local Similarity 100.0%; Pred. No. 2.5e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSNCDCLNGTGVSNKYFSNIHWCNCPKFKGGQHCBEIDKSKTCYEGNGHFYRG 60
Db 20 SNELHQPNSNCDCLNGTGVSNKYFSNIHWCNCPKFKGGQHCBEIDKSKTCYEGNGHFYRG 79

Qy 61 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKINYCRNPNRRRPWCYVOVGLK 120
Db 80 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKINYCRNPNRRRPWCYVOVGLK 139

Qy 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 140 PLVQECMVHDCADGKKPSSPPEE 162

RESULT 34
AAW24578
ID AAW24578 standard; protein; 430 AA.
XX
AC AAW24578;
XX
DT 25-MAR-2003 (revised)
DT 11-NOV-1997 (first entry)
XX
DE Inhibitor resistant urokinase.
XX
XX Urokinase; inhibitor resistant; plasminogen; human; whey acid protein;
XX plasminogen activator inhibitor-1; plasminogen activator; protease; WAP;
XX plasmin; proteolytic enzyme; clot lysis; heart attack; heart muscle.
OS Homo sapiens.
OS Synthetic.

Key Location/Qualifiers
FH Peptide 1..19
FT /note= "whey acid protein signal peptide"
FT Protein 20..430
FT /note= "urokinase"
FT Domain 20..64
FT /note= "E-domain"
FT Domain 61..150
FT /note= "Kringle-1 domain"
FT Domain 179..430
FT /note= "P-domain"
FT Misc-difference 198..203
FT /note= "deleted in modified urokinase of the invention"
XX
XX US5648253-A.
XX
XX 15-JUL-1997.
XX
XX 08-SEP-1992; 92US-00942157.
XX
XX 20-DEC-1990; 90US-00631673.
XX (TSIT-) TSI CORP.
XX Wei C;
XX WPI; 1997-372062/34.
XX N-PSDB; AAT80075.
XX Deletion-modified urokinase protein - with increased resistance to
XX inhibition by plasminogen activator inhibitor-1.
XX Disclosure; Col 15-18; 16pp; English.
XX

This sequence represents the full length urokinase, including the whey acid protein (WAP) signal peptide. This sequence has residues 179-184 of the urokinase sequence deleted to create the modified urokinase of the invention. The modified urokinase (see AAW24579) cleaves plasminogen, and has a lower binding affinity for plasminogen activator inhibitor-1 than the corresponding unmodified urokinase. Urokinase is one of two types of mammalian plasminogen activators (PA), the other being tissue type PA. PAs catalyze the conversion of the circulating zymogen plasminogen to the broad spectrum protease plasmin by limited proteolysis. The modified urokinase can be used for clot lysis, specifically to dissolve heart attack-causing clots before they cause permanent damage to heart muscle. The urokinase mutant is more resistant to inhibition by plasminogen activator inhibitors than the unmodified urokinase. It can be selectively expressed and secreted from the mammary glands of transgenic animals. (Updated on 25-MAR-2003 to correct PF field.)

Sequence 430 AA;
Query Match 100.0%; Score 837; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 2.5e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPSPNCDCINGTGVSNKYFNSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFGYRG 60
DB 20 SNELHQPSPNCDCINGTGVSNKYFNSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFGYRG 79
QY 61 KASTDTMGPRCLPWNATVLOQTYHAHRSALQGLGKKNYCRNPNRRRPWCYVQVGLK 120
DB 80 KASTDTMGPRCLPWNATVLOQTYHAHRSALQGLGKKNYCRNPNRRRPWCYVQVGLK 139
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 140 PLVQECMVHDCADGKKPSPPEE 162

RE UT 35
AA 0114
ID AAP50114 standard; protein; 431 AA.
AC AAP50114;
XX
DT 27-SEP-1991 (first entry)
DE
DE Sequence encoded by the signal sequence and noncoding region of the pro-
DE UK structural gene (Sequence II).
XX Enzyme; thrombosis therapy; embolic disease; single-chain pro-urokinase.
XX Homo sapiens.
XX

Key Location/Qualifiers
FT Peptide 1..20 /label= signal peptide
FT Domain 21..177 /label= A chain
FT Domain 179..431 /label= B chain
XX
XX EP154272-A.
XX
XX 11-SEP-1985.
XX
XX 23-FEB-1985; 85EP-00102031.
XX
XX 27-FEB-1984; 84JP-00037119.
XX 31-JAN-1985; 85JP-00017969.
XX
XX (GREC) GREEN; CROSS CORP.
XX Hiramatsu R, Kaneda T, Nagai M, Rimura H, Nishida M, Suyama T;
PI WPI; 1985-224693/37.
XX N-PSDB; AAN50138.
DR

XX Glycosylated single-chain pro-urokinase - prep'd. by cultivating animal
PT cells transformed by DNA prep'd. from m RNA.
XX
XX Disclosure; Page 8-10; 64pp; English.
XX
XX The inventors claim a method of producing single-chain pro-urokinase by
XX using as template, mRNA obtd. from cells of an established human kidney-
XX derived cell line. The urokinase is used to treat thrombosis and embolic
XX diseases as well as in the treatment of diseases in combination with
XX anticancer agents
XX
SQ Sequence 431 AA;

Query Match 100.0%; Score 837; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.6e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPSPNCDCINGTGVSNKYFNSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFGYRG 60
DB 21 SNELHQPSPNCDCINGTGVSNKYFNSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFGYRG 80
QY 61 KASTDTMGPRCLPWNATVLOQTYHAHRSALQGLGKKNYCRNPNRRRPWCYVQVGLK 120
DB 81 KASTDTMGPRCLPWNATVLOQTYHAHRSALQGLGKKNYCRNPNRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 141 PLVQECMVHDCADGKKPSPPEE 163

RESULT 36
AAP60783
ID AAP60783 standard; protein; 431 AA.
XX
AC AAP60783;
XX
XX 25-MAR-2003 (revised)
DT 23-OCT-1991 (first entry)
XX
XX Human urokinase.
XX
XX E.coli; high molecular urokinase.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Protein 21..431
XX JP61181377-A.
XX
XX 14-AUG-1986.
XX
XX 25-JAN-1985; 85JP-00011032.
XX
XX 25-JAN-1985; 85JP-00011032.
XX
XX (NISC) NISSAN CHEM IND LTD.
XX (HODO) HODOGAYA CHEM IND CO LTD.
XX (SAGA) SAGAMI CHEM RES CENTRE.
XX (CENG) CENTRAL GLASS CO LTD.
XX (NIPS) NIPPON SODA CO.
XX (TOIY) TOYO SODA MFG CO LTD.
XX
XX WPI; 1986-254744/39.
XX N-PSDB; AAN60703.
XX
XX Human urokinase gene - has N-end of aminoacid sequence coded by codon
XX used in Escherichia coli.
XX
XX Disclosure; Fig 2; 19pp; Japanese.
XX
XX The claimed gene product may be expressed in a transformed E.coli host,
XX

for the efficient production of high molecular human urokinase. The N-terminal of the protein expressed by the transforming plasmid is replaced with a codon frequently used in E.coli. (Updated on 25-MAR-2003 to correct PA field.)

Sequence 431 AA;

Query Match 100.0%; Score 837; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.6e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SNEHQVPSNCDCLNGTGVSNKYFNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 60
21 SNEHQVPSNCDCLNGTGVSNKYFNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 80

61 KASTDTMGRPCLPWNSATVLQQTTHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
81 KASTDTMGRPCLPWNSATVLQQTTHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 140

121 PLVQECMVHDCADGKPPSPPEE 143
141 PLVQECMVHDCADGKPPSPPEE 163

JLT 37

70258
AAP70258 standard; protein; 431 AA.

AAP70258;

25-MAR-2003 (revised)

19-MAY-1991 (first entry)

Sequence of human prourokinase and leader.

Cardiovascular disease treatment; fibrin affinity; thrombolytic; enzyme; protease.

Homo sapiens.

Key Location/Qualifiers
Peptide 1..20
Protein 21..431
/label= leader
/label= prourokinase

EP231883-A.

12-AUG-1987.

29-JAN-1987; 87EP-00101209.

31-JAN-1986; 86JP-00017734.

30-JAN-1987; 87JP-00018626.

(SAGA) SAGAMI CHEM RES CENTRE.
(NIPS) NIPPON SODA CO.
(CENG) CENTRAL GLASS CO LTD.
(TOYO) TOYO SODA MFG CO LTD.
(NISC) NISSAN CHEM IND LTD.
(NISC) NISSAN CHEMICAL INDS KK.

Tagawa M, Wada M, Yamada M, Yokoyama M, Numao N;

WPI; 1987-222882/32.

N-PSDB; AAN70390.

Hybrid plasminogen activator-like polypeptide - having a region for affinity to fibrin from tissue plasminogen activator and a region from prourokinase.

Disclosure; Fig 2(1-5); 64pp; English.

CC The TPA portion of the claimed hybrid polypeptide (see PT) may consist of 2 kringles from N-terminal first serine to 219th glycine of human TPA, 1 kringles from 128th serine to 219th glycine of human TPA or half a kringles from 161st methionine to 219th glycine (see AAP70257). The C-terminal half of the hybrid polypeptide may contain an AA sequence from 150th glutamine to C-terminal 411th leucine of prourokinase (see AAP70258). (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 431 AA;

Query Match 100.0%; Score 837; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.6e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGTGVSNKYFNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 60

Db 21 SNEHQVPSNCDCLNGTGVSNKYFNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWNSATVLQQTTHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120

Db 81 KASTDTMGRPCLPWNSATVLQQTTHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGKPPSPPEE 143

Db 141 PLVQECMVHDCADGKPPSPPEE 163

RESULT 38

AAP71491

ID AAP71491 standard; protein; 431 AA.

XX AAP71491;

AC AAP71491;

XX 31-OCT-2002 (revised)

DT 15-MAY-1991 (first entry)

XX Modified prourokinase.

XX Fibrin selectivity.

XX Unidentified.

XX Key Location/Qualifiers

PH Protein 21..431

FT /label= Mature product

XX EP236040-A.

PN 09-SEP-1987.

XX 24-FEB-1987; 87EP-00301567.

XX 26-FEB-1986; 86US-00833179.

PR 19-FEB-1987; 87US-00012023.

XX (COLB) COLLABORATIVE RES INC.

XX Vovis GF, Mao JI;

XX WPI; 1987-251560/36.

XX Amino acid modified form of prourokinase - having high fibrin selectivity and resistant to cleavage to the two-chain form.

XX Claim 10; Fig 1; 18pp; Japanese.

PS The modified prourokinase product has a single polypeptide chain, high fibrin selectivity and resistance to cleavage into the two-chain form.

CC The modification is at the codon for Lys158. (Updated on 31-OCT-2002 to add missing OS field.)

XX Sequence 431 AA;

Query Match 100.0%; Score 837; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.6e-55;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFFYRG 60
 DE 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFFYRG 80
 QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKKNYCRNPNRRRPPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKKNYCRNPNRRRPPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGKPPSPPEE 143
 DE 141 PLVQECMVHDCADGKPPSPPEE 163

RE UT 39
 AA 1663
 ID AAP71663 standard; protein; 431 AA.
 AC AAP71663;
 XX

DT 30-APR-1991 (first entry)

DE Modified prourokinase.

KW Prourokinase; fibrin; PUK.

OS Homo sapiens.

FH Key Location/Qualifiers
 FT Peptide 1..20 /label= sig_peptide
 FT Protein 21..431 /label= mat_protein
 FT Misc-difference 178..178 /note= "amino acid other than LYS"
 FT

PN EP236040-A.

XX 09-SEP-1987.

XX 24-FEB-1987; 87EP-00301567.

XX 26-FEB-1986; 86US-00833179.

XX 19-FEB-1987; 87US-00012023.

XX (COLB) COLLABORATIVE RES INC.

XX Vovis GF, Mao JI;

XX WPI; 1987-251560/36.

XX N-PSDB; AAN71338.

XX Amino acid modified form of prourokinase - having high fibrin selectively
 and resistant to cleavage to the two-chain form.

XX Disclosure; Fig 1; 18pp; Japanese.

XX The amino acid LYS at position 158 of the mature protein (see tag b) is
 replaced by another amino acid, achieved by site-directed mutagenesis or
 by chemical modification. The modified PUK has a single polypeptide chain,
 high fibrin activity and is resistant to cleavage to the two-chain form
 to an extent greater than its original unmodified form. The therapeutic
 specificity of single-chain PUK can be increased without decreasing
 fibrinolytic activity. The single chain form can have high fibrin
 activity and thus does not cause general bleeding in the body but rather
 can be localised to the site of the blood clot directly. See also
 AAN71368-69

XX Sequence 431 AA;

Query Match 100.0%; Score 837; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.6e-55;

Query Match 100.0%; Score 837; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.6e-55;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFFYRG 60
 DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFFYRG 80
 QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKKNYCRNPNRRRPPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKKNYCRNPNRRRPPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGKPPSPPEE 143
 DB 141 PLVQECMVHDCADGKPPSPPEE 163

RESULT 40

AAP71698

ID AAP71698 standard; protein; 431 AA.

XX AC AAP71698;

XX 30-APR-1991 (first entry)

XX Modified prourokinase (MET/SER).

XX Prourokinase; fibrin; PUK.

XX Homo sapiens.

FH Key Location/Qualifiers
 FT Peptide 1..20 /label= sig_peptide
 FT Protein 21..431 /label= mat_protein
 FT

XX EP236040-A.

XX 09-SEP-1987.

XX 24-FEB-1987; 87EP-00301567.

XX 26-FEB-1986; 86US-00833179.

XX 19-FEB-1987; 87US-00012023.

XX (COLB) COLLABORATIVE RES INC.

XX Vovis GF, Mao JI;

XX WPI; 1987-251560/36.

XX N-PSDB; AAN71368.

XX Amino acid modified form of prourokinase - having high fibrin selectively
 and resistant to cleavage to the two-chain form.

XX Disclosure; Fig 1; 18pp; Japanese.

XX The amino acid LYS at position 158 of the mature protein is replaced by
 MET and the amino acid ILE at position 160 is replaced by SER, achieved
 by site-directed mutagenesis or by chemical modification. The modified PUK
 has a single polypeptide chain, high fibrin activity and is resistant to
 cleavage to the two-chain form to an extent greater than its original
 unmodified form. The therapeutic specificity of single-chain PUK can be
 increased without decreasing fibrinolytic activity. The single chain form
 can have high fibrin activity and thus does not cause general bleeding in
 the body but rather can be localised to the site of the blood clot
 directly. See also AAN71368-69 and AAN71338

XX Sequence 431 AA;

atches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 SNELHQPNSCDCLNGTGVSNKYPNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
 D 21 SNELHQPNSCDCLNGTGVSNKYPNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 80
 Q 61 KASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
 D 81 KASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 140
 Q 121 PLVQECMVHDCADGKKPSSPPEE 143
 D 141 PLVQECMVHDCADGKKPSSPPEE 163

F JLT 41
 A /1699
 I AAP71699 standard; protein; 431 AA.
 X AAP71699;

30-APR-1991 (first entry)

Modified prourokinase (ALA).

Prourokinase; fibrin; PUK.

Homo sapiens.

Key Location/Qualifiers
 Peptide : 1..20
 Protein /label= sig_peptide
 /label= mat_protein

EP236040-A.

09-SEP-1987.

24-FEB-1987; 87EP-00301567.

26-FEB-1986; 86US-00833179.

19-FEB-1987; 87US-00012023.

(COLB) COLLABORATIVE RES INC.

Vovis GF, Mao JI;

WPI; 1987-251560/36.

N-PSDB; AAN71369.

Amino acid modified form of prourokinase - having high fibrin selectively and resistant to cleavage to the two-chain form.

Disclosure; Fig 1; 18pp; Japanese.

The amino acid LYS at position 158 of the mature protein (bases 610-612) is replaced by another amino acid (ALA), achieved by site-directed mutagenesis. The modified PUK has a single polypeptide chain, high fibrin activity and is resistant to cleavage to the two-chain form to an extent greater than its original unmodified form. The therapeutic specificity of single-chain PUK can be increased without decreasing fibrinolytic activity. The single chain form can have high fibrin activity and thus does not cause general bleeding in the body but rather can be localised to the site of the blood clot directly. See also AAN71368-69

Sequence 431 AA;

Very Match 100.0%; Score 837; DB 1; Length 431;

Best Local Similarity 100.0%; Pred. No. 2.6e-55;

atches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 SNELHQPNSCDCLNGTGVSNKYPNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60

Db 21 SNELHQPNSCDCLNGTGVSNKYPNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 80
 QY 61 KASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
 Db 81 KASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGKKPSSPPEE 143
 Db 141 PLVQECMVHDCADGKKPSSPPEE 163

RESULT 42

AAP80430

ID AAP80430 standard; protein; 431 AA.

XX AAP80430;

DT 25-MAR-2003 (revised)

DT 14-SEP-1990 (first entry)

DE Deduced AA sequence of the single chain urokinase plasminogen activator (SCU-PA) cDNA insert prepared from human Hep3 cells.

XX Single chain urokinase plasminogen activator (SCU-PA); human Hep3 cells; glyceroldehyde-3-phosphate dehydrogenase gene; thrombosis prevention; thrombosis treatment.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..20

FT Protein 21..411

XX EP288435-A.

XX 26-OCT-1988.

XX 11-APR-1988; 88EP-00810234.

XX 15-APR-1987; 87GB-00009081.

XX 16-JUN-1987; 87GB-00014059.

XX 04-DEC-1987; 87IE-00003299.

XX (CIBA) CIBA GEIGY AG.

XX Meyhack B, Heim J, Burgi R;

XX WPI; 1988-301440/43.

XX N-PSDB; AAN80981.

PT Prodn. of human single chain urokinase-type plasminogen activator - by culturing yeast strain transformed with hybrid vector contg. yeast expression control sequences.

PS Example 1; Fig 2; 48pp; English.

XX The patent is for the prodn. of human single chain urokinase-type plasminogen activator (UTPA). Mutants of scu-PA are especially those which render the protein protease resistant. Such scu-PA mutants are covalently modified at sites of proteolysis by proteases occurring in blood such as thrombin or plasmin, so that they are no longer susceptible to protease hydrolysis at these locations. The target sites include Lys135 to Lys136 (cleavage at this site generates the so-called low molecular weight form of scu-PA or LUK), Arg156 to Phe157 (susceptible to thrombin attack) and Lys158 to Ile159 (cleavage at this site by plasmin generates tuc-PA). Suitable scu-PA mutants have site specific substitutions, insertions or deletions of residues at one or more of these target sites. Especially preferred are those mutants in which one amino acid residue or both amino acid residues forming the target sites are deleted or in which at least one of these amino acid residues is replaced by another amino acid residue so that the resulting mutants are resistant to proteolytic attack. The UTPA proteins exhibit the biological

CC activity of natural human UTPA without any refolding procedure being
CC necessary. They can be used as for known PAs in humans for the prevention
CC or treatment of thrombosis or other conditions where it is desired to
CC produce local fibrinolytic or proteolytic activity. (Updated on 25-MAR-
CC 2003 to correct PR field.)
XX
SQ Sequence 431 AA;

Query Match 100.0%; Score 837; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.6e-55; Indels 0; Gaps 0;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFFYRG 60
DE 21 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFFYRG 80
QY 61 KASDTMTGRCLPWSNATVLOQTYHAHRSALQLGLGKHNCRNPNRRRPMCYVQVGLK 120
DB 81 KASDTMTGRCLPWSNATVLOQTYHAHRSALQLGLGKHNCRNPNRRRPMCYVQVGLK 140
QY 121 PLVQECWVHDCADGKKPSPPEE 143
DB 141 PLVQECWVHDCADGKKPSPPEE 163

RESULT 44
AAP91886
ID AAP91886 standard; protein; 431 AA.
XX
AC AAP91886;
XX
DT 25-MAR-2003 (revised)
DT 31-OCT-2002 (revised)
DT 16-APR-1990 (first entry)
XX
DE Pro-urokinase with signal sequence.
XX
KW pro-urokinase (pro-UK); plasminogen activator; pUK33; ss.
XX
OS Homo sapiens.

Key Location/Qualifiers
Peptide 1..20
Protein 21..431
/label= signal peptide
/label= pro-urokinase
EP265874-A.
XX
04-MAY-1988.
XX
23-OCT-1987; 87EP-00115600.
XX
23-OCT-1986; 86JP-00253078.
XX
(GREG) GREEN CROSS CORP.
XX
Amatsuji Y, Okabayashi K, Nagai M, Arimura H, Suyama T;
XX
WPI; 1988-121000/18.
XX
N-PSDB; AAN81558.
XX
Glycosylated single-chain pro-urokinase prodn. - by cultivating DHFR gene
XX
-deficient CHO-K1 cells transformed with a plasmid contg. CDNA, SV40
XX
promoter and DHFR gene.
XX
Disclosure; Page ?; 19pp; English.

CC The Arg at position 2 is encoded by TGA(sic). Possible error in the
CC specification. Should read CGA? The pro-UK gene was derived from plasmid
CC pUK33. The cDNA was synthesised using urokinase mRNA isolated from a
CC human kidney cell line. Pro-UK was cloned into a SV40 promoter-contg.
CC plasmid, down- stream of the promoter. This plasmid was then ligated to
CC a DHFR- gene contg. plasmid so that pro-UK and DHFR are inserted in

CC opposite directions. The recombinant plasmid was used to transform CHO-K1
CC cell derived DHFR gene-deficient host cells to produce glycosylated
CC single-chain pro-UK. (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 431 AA;

Query Match 100.0%; Score 837; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.6e-55; Indels 0; Gaps 0;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFFYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFFYRG 80
QY 61 KASDTMTGRCLPWSNATVLOQTYHAHRSALQLGLGKHNCRNPNRRRPMCYVQVGLK 120
DB 81 KASDTMTGRCLPWSNATVLOQTYHAHRSALQLGLGKHNCRNPNRRRPMCYVQVGLK 140
QY 121 PLVQECWVHDCADGKKPSPPEE 143
DB 141 PLVQECWVHDCADGKKPSPPEE 163

RESULT 44
AAP91886
ID AAP91886 standard; protein; 431 AA.
XX
AC AAP91886;
XX
DT 25-MAR-2003 (revised)
DT 31-OCT-2002 (revised)
DT 16-APR-1990 (first entry)
XX
DE Sequence of pro-urokinase.
XX
KW Low mol. wt. plasminogen activator; pro-urokinase; dissolve blood clot.
XX
OS Homo sapiens.

Key Location/Qualifiers
Protein 21..431
Region 170..179
/note= "Mature pro-urokinase."
/note= "Preferred initiation region for the low mol. wt. plasminogen activators."
EP316068-A.
XX
17-MAY-1989.
XX
07-OCT-1988; 88EP-00309417.
XX
09-OCT-1987; 87US-00107370.
XX
27-SEP-1988; 88US-00248727.
XX
(COLB) COLLABORATIVE RES INC.
XX
Mao JI;
XX
WPI; 1989-146601/20.
XX
N-PSDB; AAN91740.
XX
Modified low mol. wt. plasminogen activator - formed of amino acids
XX
comprising the amino acid portion of pro-urokinase from 150 to 411.

Fig 1; Page -; 27pp; English.
XX
A low mol. wt. polypeptide plasminogen activator (PA) is claimed which is
XX
formed from amino acids 150-411 of pro-urokinase. The preferred initiation
XX
region for the low mol. wt. PA is indicated (see FT). A low mol wt. PA
XX
can be injected into blood in the body in vivo to dissolve clots without
XX
harm. (Updated on 31-OCT-2002 to add missing OS field.) (Updated on 25-
XX
MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI

C field.)

X Sequence 431 AA;

S Query Match 100.0%; Score 837; DB 1; Length 431;

Best Local Similarity 100.0%; Pred. No. 2.6e-55;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHFYRG 60

D 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHFYRG 80

Q 61 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRREPCWYVQVGLK 120

D 81 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRREPCWYVQVGLK 140

Q 121 PLVQECMVHDCADGKXPSPPEE 143

D 141 PLVQECMVHDCADGKXPSPPEE 163

R JLT 45

A 92119

I AAP92119 standard; protein; 431 AA.

X AAP92119;

A 25-MAR-2003 (revised)

D 29-JUN-1990 (first entry)

X Natural human prourokinase.

D Human prourokinase; antithrombotic; derivative.

X Homo sapiens

X Key Location/Qualifiers

F Misc-difference 1 /note= "Optional in new deriv."

F Misc-difference 2 /note= "Incorporated into new deriv."

F Misc-difference 135 /note= "May be replaced by a non-basic AA in new deriv."

F Misc-difference 156 /note= "Undefined residue in new deriv."

F Misc-difference 157 /note= "Pro, Gly, Ala or Val in new deriv."

F Misc-difference 158 /note= "Lys or Arg in new deriv."

F WO8901513-A.

X 23-FEB-1989.

X 18-AUG-1988; 88WO-JP000815.

X 19-AUG-1987; 87JP-00204149.

X (SAGA) SAGAMI CHEM RES CENTRE.

X (CENG) CENTRAL GLASS CO LTD.

X (HODO) HODOGAYA CHEM KK.

X (NIPS) NIPPON SODA CO.

X (NISC) NISSAN CHEM IND LTD.

X Kobayashi Y.; Omori M, Yamada C;

X WPI; 1989-068869/09.

D N-PSDB; AAN91075.

X Antithrombotic fast-acting pro-urokinase deriv. - produced by culture of

X E coli transformant contg. new plasmid of PMUT9Q family.

X Disclosure; Fig 1; 75pp; Japanese.

XX A human prourokinase (PU) deriv. is new which is based upon residues 2-155 of natural human prourokinase. The new deriv. is produced by E. coli J103/PMUT9Q-RPK in culture. It is a fast-acting drug for the treatment CC and prevention of thrombosis. (Updated on 25-MAR-2003 to correct PA CC field.) (Updated on 25-MAR-2003 to correct PI field.)

XX SQ Sequence 431 AA;

Query Match 100.0%; Score 837; DB 1; Length 431;

Best Local Similarity 100.0%; Pred. No. 2.6e-55;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHFYRG 60

DB 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRREPCWYVQVGLK 120

DB 81 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRREPCWYVQVGLK 140

QY 121 PLVQECMVHDCADGKXPSPPEE 143

DB 141 PLVQECMVHDCADGKXPSPPEE 163

RESULT 46

AAP94764

ID AAP94764 standard; protein; 431 AA.

XX AAP94764;

XX 25-MAR-2003 (revised)

DT 27-JUN-1990 (first entry)

XX Non-glycosylated prourokinase.

DE Prourokinase; CGE 195; plasminogen activator; blood clot lysis.

XX Homo sapiens.

OS Key Location/Qualifiers

FH Region 1..20 /label= signal sequence

FT

XX EF299706-A.

XX PD 18-JAN-1989.

XX PF 12-JUL-1988; 88EP-00306334.

XX PR 13-JUL-1987; 87US-00072426.

PR 29-JUN-1988; 88US-00211279.

XX (COLB) COLLABORATIVE RES INC.

XX Baltimore D, Moir DT, Broeze RJ;

XX WPI; 1989-017204/03.

DR N-PSDB; AAN93079.

XX New non glycosylated, secreted plasminogen activator - pref. with PT asparagine replaced or deleted, useful for treating blood clots, PT expressed in non mammalian cells.

XX Disclosure; Page; 26pp; English.

XX myocardial infarction. DNA encoding the protein was sequenced from CC plasmid pCGS195, a subclone of two inserts isolated by screening a cDNA CC library prep. from kidney cell RNA. One of the original inserts, clone CC CGF31 (tag c) started in the middle of the signal sequence. Mutants of CC the sequence, pref. in which gcc (Ala) replaces aat (Aan) at nucleotides CC 1002-1004 (residue 302) are used to transform hosts for the prodn. of non

CC -glycosylated, pro-urokinase. (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 431 AA;

SC Query Match 100.0%; Score 837; DB 1; Length 431;

Best Local Similarity 100.0%; Pred. No. 2.6e-55;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFRG 60

DB 21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFRG 80

QY 61 KASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120

DB 81 KASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140

QY 121 PLVQECWVHDCADGKKPSPPEE 143

DB 141 PLVQECWVHDCADGKKPSPPEE 163

RE BT 47

AA 112

ID AAR07112 standard; protein; 431 AA.

XX AAR07112;

AC AAR07112;

XX 25-MAR-2003 (revised)

DT 24-JAN-1991 (first entry)

XX Human pro-Urokinase encoded by plasmid pUK1.

DE pro-Urokinase; transgenic mice.

XX Synthetic.

OS EP390592-A.

PN 03-OCT-1990.

XX 30-MAR-1990; 90EP-00303445.

XX 31-MAR-1989; 89JP-00078574.

XX (KYOW) KYOWA-HAKKO KOGYO KK.

PA (EXPE-) CENT INST EXPER ANIMALS.

PA (JIKK-) JIKKEN DOBUTSU CHUO KENK.

XX Sekine S, Ito S, Katsuki M;

XX WPI; 1990-299492/40.

DE N-PSDB; AAQ06049.

PT Prodn. of recombinant protein, esp. human pro-urokinase - from milk of

PT transgenic animals using promoter of bovine alpha S1 casein chromosomal

PT gene.

XX Example; Table 1; 55pp; English.

PS E.coli strain C600SF8 was transformed with recombinant plasmid containing

XX ds DNA derived from human pharynx cancer cell strain Detroit 562. 10000

CC colonies were screened and one positive clone was identified. Plasmid

CC pUK1 was isolated and found to contain the coding region and 3' non-

CC coding region of pro-UK downstream of Cys(41). Four silent substitutions

CC were identified c.f. Holmes et al., Biotechnology, vol.3, p.923 (1985) as

CC follows: (sic) Asn(254), AAC to AAT; Leu(340), CTA to CTG; Pro(345), CCC

CC to CCA; Gln(346), CAA to CAG. See also AAQ06045-Q06048 and AAQ06392.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 431 AA;

SC Query Match 100.0%; Score 837; DB 2; Length 431;

Best Local Similarity 100.0%; Pred. No. 2.6e-55;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFRG 60

DB 21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFRG 80

QY 61 KASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120

DB 81 KASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140

QY 121 PLVQECWVHDCADGKKPSPPEE 143

DB 141 PLVQECWVHDCADGKKPSPPEE 163

RESULT 48

AAR04253

ID AAR04253 standard; protein; 431 AA.

XX AAR04253;

AC AAR04253;

XX 25-MAR-2003 (revised)

DT 12-SEP-1990 (first entry)

XX Human pro-urokinase from the cDNA of clone pcUK176.

DE Human pro-urokinase; pro-urokinase; E. coli; Ptip promoter; MS-2 RBS.

XX Non-glycosylated; pro-urokinase; E. coli; Ptip promoter; MS-2 RBS.

XX Synthetic.

OS EP365894-A.

PN 02-MAY-1990.

XX 06-OCT-1989; 89EP-00118586.

XX 11-OCT-1988; 88GB-00023833.

XX (FARM) FARMITALIA ERBA SPA CARLO.

XX Brandazza A, Sarmientos P, Orsini G;

XX WPI; 1990-133447/18.

DR N-PSDB; AAQ04107.

XX Non-glycosylated pro-urokinase prodn. - using E coli B strains and E coli

PT promoter Ptip and Shine-Dalgarno sequence MS-2.

XX Disclosure; Page ?; -pp; English.

XX SER residue at position 21 is the start of the mature proUK. Non-

CC glycosylated proUK (MW 45KD) produced by E.coli B strain containing the

CC sequence. See also AAQ04101-07. (Updated on 25-MAR-2003 to correct PA

CC field.)

XX Sequence 431 AA;

SC Query Match 100.0%; Score 837; DB 2; Length 431;

Best Local Similarity 100.0%; Pred. No. 2.6e-55;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFRG 60

DB 21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFRG 80

QY 61 KASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120

DB 81 KASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140

QY 121 PLVQECWVHDCADGKKPSPPEE 143

DB 141 PLVQECWVHDCADGKKPSPPEE 163

```

F ULT 49
F 63141
F AAR63141 standard; protein; 431 AA.
F AAR63141;
F
F 25-MAR-2003 (revised)
F 09-JUN-1995 (first entry)
F
F Full length human urokinase protein.
F
F Human urokinase glycoproteins; cardiovascular diseases;
F pulmonary embolism.
F
F Homo sapiens;
F
F Key Location/Qualifiers
F Sig_peptide 1..20
F Disulfide-bond 70..151
F Disulfide-bond 91..133
F Disulfide-bond 122..146
F Disulfide-bond 168..299
F Cleavage-site 179..180
F /note= "cleavage of this site produces a bioactive two
F chain form of urokinase"
F
F Disulfide-bond 209..225
F Disulfide-bond 217..288
F Disulfide-bond 313..382
F Disulfide-bond 345..361
F Disulfide-bond 372..400
F
F EP620279-A1.
F
F 19-OCT-1994.
F
F 14-APR-1983; 94EP-00104777.
F
F 15-APR-1982; 82US-00368773.
F 14-MAR-1983; 82US-00474930.
F 14-APR-1983; 83EP-00103629.
F
F (GETH ) GENENTECH INC.
F
F Heyneker HL.; Holmes WE, Vehar GA;
F
F WPI; 1994-318362/40.
F N-PSDB; AAQ73483.
F
F Prodn. of human urokinase glycoproteins - using a recombinant expression
F system used for the treatment of vascular diseases or conditions.
F
F Claim 1; Fig 4; 41pp; English.
F
F AAQ73483 is the cDNA sequence which encodes AAR63141 the full length
F 54000 dalton human urokinase (UK) protein. This cDNA was used in the
F construction of a plasmid capable of transforming either yeast or
F vertebrate cells, enabling them to produce the 54000 dalton human UK
F protein. The UK glycoprotein produced could then be used in the treatment
F of cardiovascular diseases, including pulmonary embolism. The UK produced
F using this method had the advantage of a specific activity towards fibrin
F and exant thrombi, not demonstrated previously with UK isolated from
F natural sources. (Updated on 25-MAR-2003 to correct FN field.) (Updated
F on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct
F PR field.)
F
F Sequence 431 AA;
F
F Query Match 100.0%; Score 837; DB 2; Length 431;
F Best Local Similarity 100.0%; Pred. No. 2.6e-55;
F Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
F
F 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60
F
F 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 80
F
F 61 KASTDTMGPRPCLPWN SATVLQOTYHAHRSDALQLGLGKKNYCRNPDNRRPWCYVQVGLK 120
F
F 81 KASTDTMGPRPCLPWN SATVLQOTYHAHRSDALQLGLGKKNYCRNPDNRRPWCYVQVGLK 140
F
F 121 PLVQECMWHDCADGKKPSSPPEE 143
F
F 141 PLVQECMWHDCADGKKPSSPPEE 163
F
F Search completed: May 25, 2004, 14:53:01
F Job time : 36.9657 secs
F
F 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60

```

```

Db 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 80
Qy 61 KASTDTMGPRPCLPWN SATVLQOTYHAHRSDALQLGLGKKNYCRNPDNRRPWCYVQVGLK 120
Db 81 KASTDTMGPRPCLPWN SATVLQOTYHAHRSDALQLGLGKKNYCRNPDNRRPWCYVQVGLK 140
Qy 121 PLVQECMWHDCADGKKPSSPPEE 143
Db 141 PLVQECMWHDCADGKKPSSPPEE 163

```

```

RESULT 50
AAR47903
ID AAR47903 standard; protein; 431 AA.
XX
XX AAR47903;
XX
XX 13-JUL-1994 (first entry)
XX
XX Pro-urokinase derivative.
XX
XX Pro-urokinase; half-life; thrombolytic; thrombosis; fibrinolytic; factor.
XX
XX Homo sapiens.
XX
XX JP05336965-A.
XX
XX 21-DEC-1993.
XX
XX 17-OCT-1991; 91JP-00269615.
XX
XX 17-OCT-1991; 91JP-00269615.
XX
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX
XX WPI; 1994-030907/04.
XX
XX N-PSDB; AAQ55772.
XX
XX Novel human pro-urokinase deriva. having long half-life - with high
XX PT thrombolytic activity, useful for treatment of thrombosis.
XX
XX Disclosure; Page 15-17; 29pp; Japanese.
XX
XX Sequences (AAQ55771-72) are pro-urokinase derivatives. The products have
XX an inserted sugar moiety having an amino acid substituted, depleted or
XX inserted variant around the thrombin cleavage site. They also have a long
XX half-life allowing them to be used in the treatment of thrombosis.
XX
XX Sequence 431 AA;

```

```

Query Match 100.0%; Score 837; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.6e-55;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 80
Qy 61 KASTDTMGPRPCLPWN SATVLQOTYHAHRSDALQLGLGKKNYCRNPDNRRPWCYVQVGLK 120
Db 81 KASTDTMGPRPCLPWN SATVLQOTYHAHRSDALQLGLGKKNYCRNPDNRRPWCYVQVGLK 140
Qy 121 PLVQECMWHDCADGKKPSSPPEE 143
Db 141 PLVQECMWHDCADGKKPSSPPEE 163

```

Search completed: May 25, 2004, 14:53:01
Job time : 36.9657 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

C protein - protein search, using sw model

R on: May 25, 2004, 14:53:05 ; Search time 26.0684 Seconds
(without alignments)
1530.046 Million cell updates/sec

T le: US-09-880-503-8

F ect score: 837

S uence: 1. SNEHLQVPSNCDLNGGTCV.....QECMVHDCADKKPSPPEE 143

S ring table: BLOSUM62

S rched: Gapop 10.0 , Gapext 0.5

T al number of hits satisfying chosen parameters: 1149313

M imum DB seq length: 0

M imum DB seq length: 2000000000

P t-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

D atabase : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US02_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09E_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

R ult No.	Score	Query Match	Length	ID	Description
1	837	100.0	143	9	US-09-880-503-8
2	837	100.0	337	14	US-10-106-698-6266
3	837	100.0	337	15	US-10-264-049-2927
4	837	100.0	411	9	US-09-880-503-3
5	837	100.0	431	9	US-09-264-486B-1
6	837	100.0	431	12	US-10-411-037-34
7	837	100.0	431	12	US-10-411-026-34
8	837	100.0	431	13	US-10-076-421-2
9	837	100.0	431	13	US-10-171-311-184
10	837	100.0	431	14	US-10-193-656-4
11	837	100.0	431	14	US-10-301-822-161
12	837	100.0	431	14	US-10-131-985-21
13	837	100.0	431	15	US-10-295-027-414
14	837	100.0	431	15	US-10-295-027-1275
15	837	100.0	431	16	US-10-410-962-34

ALIGNMENTS

RESULT 1

US-09-880-503-8

; Sequence 8, Application US/09880503

; Patent No. US20020131964A1

; GENERAL INFORMATION:

; APPLICANT: CINES, Douglas B

; APPLICANT: HIGAZI, Abd Al-Roof

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND

837	100.0	431	16	US-10-411-049-34	Sequence 34, Appl
837	100.0	437	12	US-10-087-192-594	Sequence 594, App
835	99.8	431	14	US-10-247-671-149	Sequence 149, App
827	98.8	411	15	US-10-407-821-2	Sequence 2, Appl
821	98.1	431	12	US-10-282-174-562	Sequence 562, App
815	97.4	445	15	US-10-360-101-266	Sequence 266, App
793	94.7	135	9	US-09-880-503-4	Sequence 4, Appl
793	94.7	138	9	US-09-984-186-12	Sequence 12, Appl
793	94.7	138	14	US-10-237-667-12	Sequence 12, Appl
793	94.7	138	14	US-10-237-708-12	Sequence 12, Appl
793	94.7	138	14	US-10-237-866-12	Sequence 12, Appl
793	94.7	138	14	US-10-237-871-12	Sequence 12, Appl
793	94.7	138	14	US-10-237-634-12	Sequence 12, Appl
793	94.7	138	16	US-10-702-536-12	Sequence 12, Appl
793	94.7	138	16	US-10-702-636-12	Sequence 12, Appl
793	94.7	403	9	US-09-880-503-6	Sequence 6, Appl
595.5	71.1	650	15	US-10-401-077-1	Sequence 1, Appl
569	68.0	432	12	US-10-087-192-591	Sequence 591, App
554	66.2	96	9	US-09-880-503-9	Sequence 9, Appl
510	60.9	88	9	US-09-880-503-1	Sequence 1, Appl
505	60.3	322	12	US-10-424-999-21	Sequence 21, Appl
505	60.3	322	12	US-10-425-000-41	Sequence 41, Appl
505	60.3	322	15	US-10-233-675A-21	Sequence 21, Appl
501	59.9	91	12	US-10-424-999-37	Sequence 37, Appl
501	59.9	221	12	US-10-424-999-36	Sequence 36, Appl
501	59.9	687	12	US-10-424-999-17	Sequence 17, Appl
501	59.9	687	12	US-10-425-000-37	Sequence 37, Appl
501	59.9	687	15	US-10-233-675A-17	Sequence 17, Appl
500	59.7	87	12	US-10-424-999-9	Sequence 9, Appl
500	59.7	87	12	US-10-425-000-29	Sequence 29, Appl
500	59.7	86	15	US-10-233-675A-9	Sequence 9, Appl
499	59.6	86	12	US-10-424-999-5	Sequence 5, Appl
499	59.6	86	12	US-10-424-999-62	Sequence 62, Appl
499	59.6	86	12	US-10-425-000-97	Sequence 97, Appl
499	59.6	86	15	US-10-233-675A-5	Sequence 5, Appl
498	59.5	688	12	US-10-233-675A-22	Sequence 22, Appl
498	59.5	688	12	US-10-424-999-18	Sequence 18, Appl
498	59.5	688	15	US-10-425-000-38	Sequence 38, Appl
498	59.5	688	12	US-10-233-675A-18	Sequence 18, Appl
498	59.5	689	12	US-10-425-000-33	Sequence 33, Appl
498	59.5	689	12	US-10-425-000-33	Sequence 33, Appl
496	59.3	91	12	US-10-424-999-35	Sequence 35, Appl
496	59.3	221	12	US-10-424-999-34	Sequence 34, Appl
495	59.1	87	12	US-10-424-999-10	Sequence 10, Appl
495	59.1	87	12	US-10-425-000-30	Sequence 30, Appl
495	59.1	87	15	US-10-233-675A-10	Sequence 10, Appl
495	59.1	672	12	US-10-424-999-15	Sequence 15, Appl
495	59.1	672	12	US-10-425-000-35	Sequence 35, Appl
495	59.1	672	15	US-10-233-675A-15	Sequence 15, Appl
495	59.1	674	12	US-10-424-999-14	Sequence 14, Appl
495	59.1	674	12	US-10-425-000-34	Sequence 34, Appl
495	59.1	674	15	US-10-233-675A-14	Sequence 14, Appl
494	59.0	86	12	US-10-424-999-1	Sequence 1, Appl
494	59.0	86	15	US-10-233-675A-1	Sequence 1, Appl
494	59.0	322	12	US-10-424-999-20	Sequence 20, Appl
494	59.0	322	12	US-10-425-000-40	Sequence 40, Appl
494	59.0	322	15	US-10-233-675A-20	Sequence 20, Appl
489	58.4	86	12	US-10-424-999-7	Sequence 7, Appl
489	58.4	86	15	US-10-233-675A-7	Sequence 7, Appl

```

; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; RIOR APPLICATION NUMBER: US 60/212,847
; RIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
; US 9-880-503-8

Query Match 100.0%; Score 837; DB 9; Length 143;
Best Local Similarity 100.0%; Pred. No. 2.4e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 60
Qy 61 KASTDTMGPRCLPWNATVLTQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
Db 61 KASTDTMGPRCLPWNATVLTQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
Qy 121 PLVQECWVHDCADGKKPSSPPPEE 143
Db 121 PLVQECWVHDCADGKKPSSPPPEE 143

RE IT 2
US 9-106-698-6266
; Sequence 6266, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; RIOR APPLICATION NUMBER: PCT/US00/26524
; RIOR FILING DATE: 2000-09-28
; RIOR APPLICATION NUMBER: US 60/157,137
; RIOR FILING DATE: 1999-09-29
; RIOR APPLICATION NUMBER: US 60/163,280
; RIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6266
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; US 0-106-698-6266

Query Match 100.0%; Score 837; DB 14; Length 337;
Best Local Similarity 100.0%; Pred. No. 6.1e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 60
Db 27 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 86
Qy 61 KASTDTMGPRCLPWNATVLTQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
Db 87 KASTDTMGPRCLPWNATVLTQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 146
Qy 121 PLVQECWVHDCADGKKPSSPPPEE 143
Db 121 PLVQECWVHDCADGKKPSSPPPEE 143

RE IT 3
US 0-106-698-6266
; Sequence 6266, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; RIOR APPLICATION NUMBER: PCT/US00/26524
; RIOR FILING DATE: 2000-09-28
; RIOR APPLICATION NUMBER: US 60/157,137
; RIOR FILING DATE: 1999-09-29
; RIOR APPLICATION NUMBER: US 60/163,280
; RIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6266
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; US 0-106-698-6266

Query Match 100.0%; Score 837; DB 14; Length 337;
Best Local Similarity 100.0%; Pred. No. 6.1e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 60
Db 27 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 86
Qy 61 KASTDTMGPRCLPWNATVLTQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
Db 87 KASTDTMGPRCLPWNATVLTQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 146
Qy 121 PLVQECWVHDCADGKKPSSPPPEE 143
Db 147 PLVQECWVHDCADGKKPSSPPPEE 169

```

```

US-10-264-049-2927
; Sequence 2927, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2927
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-264-049-2927

Query Match 100.0%; Score 837; DB 15; Length 337;
Best Local Similarity 100.0%; Pred. No. 6.1e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 60
Db 27 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 86
Qy 61 KASTDTMGPRCLPWNATVLTQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
Db 87 KASTDTMGPRCLPWNATVLTQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 146
Qy 121 PLVQECWVHDCADGKKPSSPPPEE 143
Db 147 PLVQECWVHDCADGKKPSSPPPEE 169

RESULT 4
US-09-880-503-3
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-880-503-3

Query Match 100.0%; Score 837; DB 9; Length 411;
Best Local Similarity 100.0%; Pred. No. 7.6e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 60
Qy 61 KASTDTMGPRCLPWNATVLTQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
Db 61 KASTDTMGPRCLPWNATVLTQTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
Qy 121 PLVQECWVHDCADGKKPSSPPPEE 143

```

Db 121 PLVQECMVHDCADGKKPSPPEE 143
RESULT 5
US-09-264-468B-1
; Sequence 1, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walzer, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Roninon
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.P1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
; OTHER INFORMATION: Leader sequence
; NAME/KEY: VARIANT
; LOCATION: (279)...(279)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (302)...(302)
; OTHER INFORMATION: Xaa = any amino acid
US-09-264-468B-1

Query Match 100.0%; Score 837; DB 9; Length 431;
Best Local Similarity 100.0%; Pred. No. 8e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCCEIDKSKTCYEGNGHFGYRG 60
Db 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCCEIDKSKTCYEGNGHFGYRG 80
QY 61 KASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
Db 81 KASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKKPSPPEE 143
Db 141 PLVQECMVHDCADGKKPSPPEE 163

RESULT 6
US-10-411-037-34
; Sequence 34, Application US/10411037
; Publication No. US20040043446A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bows, Carlyn
; TITLE OF INVENTION: ALPHA GALACTOSIDASE A: REMODELING AND GLYCOCONJUGATION OF ALPHA
; TITLE OF INVENTION: GALACTOSIDASE A

; FILE REFERENCE: 040853-01-5082
; CURRENT APPLICATION NUMBER: US/10/411,037
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-037-34
Query Match 100.0%; Score 837; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 8e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCCEIDKSKTCYEGNGHFGYRG 60
Db 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCCEIDKSKTCYEGNGHFGYRG 80
QY 61 KASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
Db 81 KASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKKPSPPEE 143
Db 141 PLVQECMVHDCADGKKPSPPEE 163

RESULT 7
US-10-411-026-34
; Sequence 34, Application US/10411026
; Publication No. US20040063911A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: PROTEIN REMODELING METHODS AND PROTEINS/PEPTIDES PRODUCED BY THE
; TITLE OF INVENTION: METHODS
; FILE REFERENCE: 040853-01-5053
; CURRENT APPLICATION NUMBER: US/10/411,026
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 34

; LENGTH: 431

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-411-026-34

Query Match

Best Local Similarity 100.0%; Score 837; DB 12; Length 431;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHFYRG 60

DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGPRCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVOVGLK 120

DB 81 KASTDTMGPRCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVOVGLK 140

QY 121 PLVQECMVHDCADGKKPSSPPEE 143

DB 141 PLVQECMVHDCADGKKPSSPPEE 163

RESULT 8

US-10-076-421-2

; Sequence 2, Application US/10076421

; Publication No. US20020193304A1

; GENERAL INFORMATION:

; APPLICANT: WADA, MANABU

; APPLICANT: WADA, NAKAO

; TITLE OF INVENTION: ANTI-HIV AGENTS

; FILE REFERENCE: HAYAK-9

; CURRENT APPLICATION NUMBER: US/10/076,421

; CURRENT FILING DATE: 2002-05-17

; PRIOR APPLICATION NUMBER: JP 2001-42655

; PRIOR FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: JP 2001-184284

; PRIOR FILING DATE: 2001-06-19

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 431

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-076-421-2

Query Match

Best Local Similarity 100.0%; Score 837; DB 13; Length 431;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHFYRG 60

DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGPRCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVOVGLK 120

DB 81 KASTDTMGPRCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVOVGLK 140

QY 121 PLVQECMVHDCADGKKPSSPPEE 143

DB 141 PLVQECMVHDCADGKKPSSPPEE 163

RESULT 9

US-10-171-311-184

; Sequence 184, Application US/10171311

; Publication No. US20030087270A1

; GENERAL INFORMATION:

; APPLICANT: Schlegel, Robert

; APPLICANT: Chen, Yan

; APPLICANT: Zhao, Xumei

; APPLICANT: Monahan, John

Query Match

Best Local Similarity 100.0%; Score 837; DB 14; Length 431;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; APPLICANT: Kamatkar, Shubhangi

; APPLICANT: Giatt, Karen

; APPLICANT: Gannavarapu, Manjula

; APPLICANT: Hoersh, Sebastian

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY

; TITLE OF INVENTION: OF CERVICAL CANCER

; FILE REFERENCE: MRI-035

; CURRENT APPLICATION NUMBER: US/10/171,311

; CURRENT FILING DATE: 2002-06-12

; PRIOR APPLICATION NUMBER: US 60/298,159

; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/298,155

; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/335,936

; PRIOR FILING DATE: 2001-11-14

; NUMBER OF SEQ ID NOS: 238

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 184

; LENGTH: 431

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-171-311-184

Query Match

Best Local Similarity 100.0%; Score 837; DB 14; Length 431;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHFYRG 60

DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGPRCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVOVGLK 120

DB 81 KASTDTMGPRCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVOVGLK 140

QY 121 PLVQECMVHDCADGKKPSSPPEE 143

DB 141 PLVQECMVHDCADGKKPSSPPEE 163

RESULT 10

US-10-193-656-4

; Sequence 4, Application US/10193656

; Publication No. US20030096733A1

; GENERAL INFORMATION:

; APPLICANT: NY, Tor

; APPLICANT: HOLMDAHL, Rikard

; APPLICANT: LI, Jinhua

; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS

; FILE REFERENCE: 3810/10577-US3

; CURRENT APPLICATION NUMBER: US/10/193,656

; CURRENT FILING DATE: 2002-07-10

; PRIOR APPLICATION NUMBER: US 60/304,461

; PRIOR FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: US 60/304,490

; PRIOR FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: US 60/305,182

; PRIOR FILING DATE: 2001-07-13

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 431

; TYPE: PRT

; ORGANISM: Homo sapiens

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: GenBank / P00749

; DATABASE ENTRY DATE: 1986-07-21

; RELEVANT RESIDUES: (1)...(431)

US-10-193-656-4

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGPRCLPWSNATVLOQTYHAHRSDALQLGLGKHNTCRNPDNRNRRPWCYVQVGLK 120
DB 81 KASTDTMGPRCLPWSNATVLOQTYHAHRSDALQLGLGKHNTCRNPDNRNRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGKKPSSPPEE 143
DB 141 PLVQECMVHDCADGKKPSSPPEE 163

RESULT 11
US-10-301-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burtgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MP001-0292RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match 100.0%; Score 837; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 8e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGPRCLPWSNATVLOQTYHAHRSDALQLGLGKHNTCRNPDNRNRRPWCYVQVGLK 120
DB 81 KASTDTMGPRCLPWSNATVLOQTYHAHRSDALQLGLGKHNTCRNPDNRNRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGKKPSSPPEE 143
DB 141 PLVQECMVHDCADGKKPSSPPEE 163

RESULT 12
US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V

APPLICANT: Huggins, Jonathan P
APPLICANT: McIntosh, Fraser S
APPLICANT: Occleston, Nicholas L
TITLE OF INVENTION: Composition
FILE REFERENCE: PCS 10391A
CURRENT APPLICATION NUMBER: US/10/131,985
CURRENT FILING DATE: 2002-04-25
PRIOR APPLICATION NUMBER: US/09/726,295
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: GB 9930768.8
PRIOR FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-985-21

Query Match 100.0%; Score 837; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 8e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGPRCLPWSNATVLOQTYHAHRSDALQLGLGKHNTCRNPDNRNRRPWCYVQVGLK 120
DB 81 KASTDTMGPRCLPWSNATVLOQTYHAHRSDALQLGLGKHNTCRNPDNRNRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGKKPSSPPEE 143
DB 141 PLVQECMVHDCADGKKPSSPPEE 163

RESULT 13
US-10-295-027-414
; Sequence 414, Application US/10295027
; Publication No. US2003023350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250

```
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 414
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-414

Query Match      100.0%; Score 837; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 8e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
Db      21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 80
QY      61 KASTDTMGRPCLPWN SATVLQQTTHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
Db      81 KASTDTMGRPCLPWN SATVLQQTTHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 140
QY      121 PLVQECWHDCAADGKKPSSPPEE 143
Db      141 PLVQECWHDCAADGKKPSSPPEE 163

RESULT 14
US-10-295-027-1275
; Sequence 1275, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1275
```

```
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1275

Query Match      100.0%; Score 837; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 8e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
Db      21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 80
QY      61 KASTDTMGRPCLPWN SATVLQQTTHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
Db      81 KASTDTMGRPCLPWN SATVLQQTTHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 140
QY      121 PLVQECWHDCAADGKKPSSPPEE 143
Db      141 PLVQECWHDCAADGKKPSSPPEE 163

RESULT 15
US-10-410-962-34
; Sequence 34, Application US/10410962
; Publication No. US20040077836A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bowe, Caryn
; TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR: REMODELING AND
; FILE REFERENCE: 040853-01-5054
; CURRENT APPLICATION NUMBER: US/10/410,962
; PRIOR FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-410-962-34

Query Match      100.0%; Score 837; DB 16; Length 431;
Best Local Similarity 100.0%; Pred. No. 8e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
Db      21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 80
QY      61 KASTDTMGRPCLPWN SATVLQQTTHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
Db      81 KASTDTMGRPCLPWN SATVLQQTTHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 140
```

QY 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 141 PLVQECMVHDCADGKKPSSPPEE 163

RESULT 16
US-10-411-049-34
; Sequence 34, Application US/10411049
; Publication No. US20040082026A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bowe, Caryn
; TITLE OF INVENTION: INTERFERON ALPHA: REMODELING AND GLYCOCONJUGATION OF INTERFERON
; FILE REFERENCE: 040853-01-5055
; CURRENT APPLICATION NUMBER: US/10/411,049
; PRIOR FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-049-34

Query Match 100.0%; Score 837; DB 16; Length 431;
Best Local Similarity 100.0%; Pred. No. 8e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRRPPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRRPPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 141 PLVQECMVHDCADGKKPSSPPEE 163

RESULT 17
US-10-087-192-594
; Sequence 594, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 594
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-594

Query Match 100.0%; Score 837; DB 12; Length 437;
Best Local Similarity 100.0%; Pred. No. 8.1e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 27 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 86

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRRPPWCYVQVGLK 120
Db 87 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRRPPWCYVQVGLK 146

QY 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 147 PLVQECMVHDCADGKKPSSPPEE 169

RESULT 18
US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match 99.8%; Score 835; DB 14; Length 431;
Best Local Similarity 99.3%; Pred. No. 1.3e-75;
Matches 142; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRRPPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRRPPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 141 PLVQECMVHDCADGKKPSSPPEE 163

RESULT 19

```
US-10-407-821-2
; Sequence 2, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
; FILE REFERENCE: UTSN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-2

Query Match 98.8%; Score 827; DB 15; Length 411;
Best Local Similarity 99.3%; Pred. No. 7,7e-75;
Matches 142; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRRRRCVYQVGLK 120
DB 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRRRRCVYQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 LLVQECMVHDCADGKKPSPPEE 143

RESULT 20
US-10-282-174-562
; Sequence 562, Application US/10282174
; Publication No. US20030224380A1
; GENERAL INFORMATION:
; APPLICANT: Becker, Kenneth David
; APPLICANT: Vellicellebi, Gonul
; APPLICANT: Elliot, Kathryn J.
; APPLICANT: Wang, Xin
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Bertram, Lars
; APPLICANT: Saunders, Aleister J.
; APPLICANT: Mullin, Kristina M.
; APPLICANT: Sampson, Andrew Johnson
; APPLICANT: Blacker, Deborah Lyne
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
; TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
; TITLE OF INVENTION: NEURODEGENERATIVE DISEASES
; FILE REFERENCE: 37481-3308
; CURRENT APPLICATION NUMBER: US/10/282,174
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US 60/339,525
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 60/338,010
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/336,929
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/338,363
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/337,052
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/368,919
; PRIOR FILING DATE: 2002-03-28
```

```
; NUMBER OF SEQ ID NOS: 564
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 562
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: 15, 58, 141, 214, 231, 274, 366
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-282-174-562

Query Match 98.1%; Score 821; DB 12; Length 431;
Best Local Similarity 98.6%; Pred. No. 3.2e-74;
Matches 141; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRRRRCVYQVGLK 120
DB 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRRRRCVYQVGLK 140
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 141 XLVQECMVHDCADGKKPSPPEE 163

RESULT 21
US-10-360-101-266
; Sequence 266, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 266
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of urokinase
US-10-360-101-266

Query Match 97.4%; Score 815; DB 15; Length 445;
Best Local Similarity 97.3%; Pred. No. 1.3e-73;
Matches 143; Conservative 0; Mismatches 0; Indels 4; Gaps 2;

QY 1 SNELHQVPSNCDLNGGTCVSNKYF--SNHWCNCPKFKGGQHCIEDKSKTCYEGNGHFY 58
DB 21 SNELHQVPSNCDLNGGTCVSNKYFPTSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFY 80
QY 59 RGKASTDTMGRCLPWN SATVLQOTY--HAHRSDALQLGLGKHNYCRNPDNRRRRCVYQ 116
DB 81 RGKASTDTMGRCLPWN SATVLQOTYFTHAHRSDALQLGLGKHNYCRNPDNRRRRCVYQ 140
QY 117 VGLKPLVQECMVHDCADGKKPSPPEE 143
DB 141 VGLKPLVQECMVHDCADGKKPSPPEE 167

RESULT 22
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
```


ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12

Query Match 94.7%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 6.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKXFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKXFGGQHCIEDKSKTCYEGNGHFYRG 63
QY 61 KASDTMTGRPCLPWSATVLQOQTYHAHRS DALQLGLGKKNYCRNPNRRPWCYVQVGLK 120
DB 64 KASDTMTGRPCLPWSATVLQOQTYHAHRS DALQLGLGKKNYCRNPNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 25
US-10-237-708-12
Sequence 12, Application US/10237708
Publication No. US20030036170A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,708
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PC7/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.

ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12

Query Match 94.7%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 6.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKXFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKXFGGQHCIEDKSKTCYEGNGHFYRG 63
QY 61 KASDTMTGRPCLPWSATVLQOQTYHAHRS DALQLGLGKKNYCRNPNRRPWCYVQVGLK 120
DB 64 KASDTMTGRPCLPWSATVLQOQTYHAHRS DALQLGLGKKNYCRNPNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 26
US-10-237-866-12
Sequence 12, Application US/10237866
Publication No. US20030036171A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PC7/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.

REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match 94.7%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 6.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGPRCLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 64 KASTDTMGPRCLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 27

US-10-237-871-12
Sequence 12, Application US/10237871
Publication No. US20030036172A1

GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard
Fournier, Alain
Guittion, Jean-Dominique
Jung, Gerard
Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/237,871
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619

REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-871-12

Query Match 94.7%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 6.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGPRCLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 64 KASTDTMGPRCLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 28

US-10-237-624-12

Sequence 12, Application US/10237624

Publication No. US20030082747A1

GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard
Fournier, Alain
Guittion, Jean-Dominique
Jung, Gerard
Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/237,624
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-624-12

Query Match          94.7%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 6.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCIDKSKTCYEGNGHFYRG 60
DB 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKNYCRNPNRRPWCYVQVGLK 120
DB 64 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKNYCRNPNRRPWCYVQVGLK 123
QY 121 PLVQECWVHDCADGK 135
DB 124 PLVQECWVHDCADGK 138

RESULT 29
US-10-702-536-12
; Sequence 12, Application US/10702536
; Publication No. US20040086976A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/702,536
; FILING DATE: 07-Nov-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
;

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-702-536-12

Query Match          94.7%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 6.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCIDKSKTCYEGNGHFYRG 60
DB 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKNYCRNPNRRPWCYVQVGLK 120
DB 64 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKNYCRNPNRRPWCYVQVGLK 123
QY 121 PLVQECWVHDCADGK 135
DB 124 PLVQECWVHDCADGK 138

RESULT 30
US-10-702-636-12
; Sequence 12, Application US/10702636
; Publication No. US20040086977A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/702,636
; FILING DATE: 06-Nov-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
;

```

```
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-702-636-12

Query Match          94.7%; Score 793; DB 16; Length 138;
Best Local Similarity 100.0%; Pred. No. 6.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHGVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60
Db 4 SNELHGVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRREPWCVYQVGLK 120
Db 64 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRREPWCVYQVGLK 123
QY 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 31
US-09-880-503-6
; Sequence 6, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-6

Query Match          94.7%; Score 793; DB 9; Length 403;
Best Local Similarity 100.0%; Pred. No. 2e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHGVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60
Db 1 SNELHGVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRREPWCVYQVGLK 120
Db 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRREPWCVYQVGLK 120
QY 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135

RESULT 32
US-10-401-077-1
; Sequence 1, Application US/10401077
; Publication No. US20040002137A1
; GENERAL INFORMATION:
; APPLICANT: Hung, Paul Porwen
; APPLICANT: Wu, Bryan T. H.
```

```
; TITLE OF INVENTION: HUMAN TISSUE UROKINASE TYPE PLASMINOGEN
; TITLE OF INVENTION: ACTIVATOR PRODUCTION
; FILE REFERENCE: 12133-006001
; CURRENT APPLICATION NUMBER: US/10/401,077
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/371,013
; PRIOR FILING DATE: 2002-04-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-401-077-1

Query Match          71.1%; Score 595.5; DB 15; Length 650;
Best Local Similarity 77.9%; Pred. No. 2.4e-51;
Matches 106; Conservative 7; Mismatches 18; Indels 5; Gaps 3;

QY 3 ELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFY 58
Db 77 QCHSVFVKSCSEPERCFNGGTCQALYFSDF-VCCQPEGAGKCCEDTATCYEGNGHFY 135
QY 59 RGKASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRREPWCVYQV 118
Db 136 RGKASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRREPWCVYQV 195
QY 119 LKPLVQECMVHDCADG 134
Db 196 LKPLVQECMVHDCSEG 211

RESULT 33
US-10-087-192-591
; Sequence 591, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 591
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-591

Query Match          68.0%; Score 569; DB 12; Length 433;
Best Local Similarity 70.4%; Pred. No. 7.2e-49;
Matches 95; Conservative 15; Mismatches 25; Indels 0; Gaps 0;

QY 9 SNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMG 68
Db 30 SNCGCQGGVCVSYKYFSRIRRCSCPRKFGQGHCEIDKSKTCYHNGNDSYRGKANTDTKG 89
QY 69 RPLCPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRREPWCVYQVGLPLVQECMV 128
Db 90 RPLCPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRREPWCVYQVGLPLVQECMV 149
QY 129 HDCADGKPSPPPEE 143
Db 150 HDCSLSKPSSSSVDQ 164
```

```
RESULT 34
US-09-880-503-9
; Sequence 9, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 9
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-9

Query Match 66.2%; Score 554; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 4.6e-48;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 107
Db 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60

QY 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 143
Db 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96

RESULT 35
US-09-880-503-1
; Sequence 1, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-1

Query Match 60.9%; Score 510; DB 9; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.1e-43;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 107
Db 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60

QY 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135
Db 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88

RESULT 36
US-10-424-999-21
; Sequence 21, Application US/10424999
```

```
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 21
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-21

Query Match 60.3%; Score 505; DB 12; Length 322;
Best Local Similarity 92.8%; Pred. No. 1.4e-42;
Matches 87; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPD 106
Db 1 AKTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPD 60

QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSP 140
Db 61 NRRRPWCYVQVGLKPLVQECMVHDCADRLPRGP 94

RESULT 37
US-10-425-000-41
; Sequence 41, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; PRIOR FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 41
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-41

Query Match 60.3%; Score 505; DB 12; Length 322;
Best Local Similarity 92.8%; Pred. No. 1.4e-42;
Matches 87; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPD 106
Db 1 AKTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPD 60

QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSP 140
Db 61 NRRRPWCYVQVGLKPLVQECMVHDCADRLPRGP 94
```

```
RESULT 38
US-10-233-675A-21
; Sequence 21, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/233,675A
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 21
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion protein human abrogen
US-10-233-675A-21
Query Match 60.3%; Score 505; DB 15; Length 322;
Best Local Similarity 92.6%; Pred. No. 1.4e-42;
Matches 87; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 47 SKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHYCRNPD 106
Db 1 AKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHYCRNPD 60
QY 107 NRRPWCYVQGLKPLVQECMVHDCADGKPPSP 140
Db 61 NRRPWCYVQGLKPLVQECMVHDCADRLPRGP 94
RESULT 39
US-10-424-999-37
; Sequence 37, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 37
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Abrogen D43
US-10-424-999-37
Query Match 59.9%; Score 501; DB 12; Length 91;
Best Local Similarity 92.6%; Pred. No. 9.2e-43;
Matches 88; Conservative 1; Mismatches 2; Indels 4; Gaps 1;
QY 39 GQCEIDKSKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGK 98
Db 1 GSH-----MAKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGK 56
QY 99 HNYCRNPDRRRPWCYVQGLKPLVQECMVHDCAD 133
Db 1 AKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHYCRNPD 106
```

```
Db 57 HNYCRNPDRRRPWCYVQGLKPLVQECMVHDCAD 91
RESULT 40
US-10-424-999-36
; Sequence 36, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 36
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: TrxA-Abrogen D43 fusion protein
US-10-424-999-36
Query Match 59.9%; Score 501; DB 12; Length 221;
Best Local Similarity 92.6%; Pred. No. 2.4e-42;
Matches 88; Conservative 1; Mismatches 2; Indels 4; Gaps 1;
QY 39 GQCEIDKSKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGK 98
Db 131 GSH-----MAKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGK 186
QY 99 HNYCRNPDRRRPWCYVQGLKPLVQECMVHDCAD 133
Db 187 HNYCRNPDRRRPWCYVQGLKPLVQECMVHDCAD 221
RESULT 41
US-10-424-999-17
; Sequence 17, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 17
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-17
Query Match 59.9%; Score 501; DB 12; Length 687;
Best Local Similarity 97.7%; Pred. No. 8.2e-42;
Matches 86; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 47 SKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHYCRNPD 106
Db 1 AKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHYCRNPD 60
```

QY 107 NRRPWCYVQVGLKPLVQECMWHDCADG 134
Db 61 NRRPWCYVQVGLKPLVQECMWHDCADG 88

RESULT 42
US-10-425-000-37
; Sequence 37, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blancher, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 37
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-37

Query Match 59.9%; Score 501; DB 12; Length 687;
Best Local Similarity 97.7%; Pred. No. 8.2e-42;
Matches 86; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGPRCLPWSATVLTQTYHAHRSDALQLGLGKHNCRNPD 106
Db 1 AKTCYEGNGHFYRGKASTDTMGPRCLPWSATVLTQTYHAHRSDALQLGLGKHNCRNPD 60
QY 107 NRRPWCYVQVGLKPLVQECMWHDCADG 134
Db 61 NRRPWCYVQVGLKPLVQECMWHDCADG 88

RESULT 43
US-10-233-675A-17
; Sequence 17, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 17
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion protein human abrogen
US-10-233-675A-17

Query Match 59.9%; Score 501; DB 15; Length 687;
Best Local Similarity 97.7%; Pred. No. 8.2e-42;
Matches 86; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGPRCLPWSATVLTQTYHAHRSDALQLGLGKHNCRNPD 106
Db 1 AKTCYEGNGHFYRGKASTDTMGPRCLPWSATVLTQTYHAHRSDALQLGLGKHNCRNPD 60
QY 107 NRRPWCYVQVGLKPLVQECMWHDCADG 134
Db 61 NRRPWCYVQVGLKPLVQECMWHDCADG 88

RESULT 44
US-10-424-999-9
; Sequence 9, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blancher, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 9
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human abrogen as secreted from pMB063 (abrogen D43)
US-10-424-999-9

Query Match 59.7%; Score 500; DB 12; Length 87;
Best Local Similarity 98.9%; Pred. No. 1.1e-42;
Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGPRCLPWSATVLTQTYHAHRSDALQLGLGKHNCRNPD 106
Db 1 AKTCYEGNGHFYRGKASTDTMGPRCLPWSATVLTQTYHAHRSDALQLGLGKHNCRNPD 60
QY 107 NRRPWCYVQVGLKPLVQECMWHDCAD 133
Db 61 NRRPWCYVQVGLKPLVQECMWHDCAD 87

RESULT 45
US-10-425-000-29
; Sequence 29, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blancher, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 29
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human abrogen (D43)
US-10-425-000-29

Query Match 59.7%; Score 500; DB 12; Length 87;

Best Local Similarity 98.9%; Pred. No. 1.1e-42;
Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKKNYCRNPD 106
:|||||
Db 1 AKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKKNYCRNPD 60

QY 107 NRRPWCYVQVGLKPLVQECMVHDCAD 133
:|||||
Db 61 NRRPWCYVQVGLKPLVQECMVHDCAD 87

RESULT 46

US-10-233-675A-9

; Sequence 9, Application US/10233675A

; Publication No. US20030228298A1

; GENERAL INFORMATION:

; APPLICANT: Nesbit, Mark

; APPLICANT: Fong, Timothy

; APPLICANT: Brockstedt, Dirk

; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for

; FILE REFERENCE: ST01027

; CURRENT APPLICATION NUMBER: US/10/233,675A

; CURRENT FILING DATE: 2002-09-04

; PRIOR FILING DATE: 2001-09-04

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 9

; LENGTH: 87

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: human derived fusion protein

US-10-233-675A-9

Query Match 59.7%; Score 500; DB 15; Length 87;
Best Local Similarity 98.9%; Pred. No. 1.1e-42;
Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKKNYCRNPD 106
:|||||
Db 1 AKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKKNYCRNPD 60

QY 107 NRRPWCYVQVGLKPLVQECMVHDCAD 133
:|||||
Db 61 NRRPWCYVQVGLKPLVQECMVHDCAD 87

RESULT 47

US-10-424-999-5

; Sequence 5, Application US/10424999

; Publication No. US20040052810A1

; GENERAL INFORMATION:

; APPLICANT: Nesbit, Mark

; APPLICANT: Cameron, Beatrice

; APPLICANT: Blanche, Francis

; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for

; FILE REFERENCE: ST01027-A

; CURRENT APPLICATION NUMBER: US/10/424,999

; CURRENT FILING DATE: 2003-04-29

; PRIOR FILING DATE: 2002-09-04

; NUMBER OF SEQ ID NOS: 70

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 5

; LENGTH: 86

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Human abrogen

US-10-424-999-5

Query Match 59.6%; Score 499; DB 12; Length 86;

Best Local Similarity 100.0%; Pred. No. 1.4e-42;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKKNYCRNPD 107
:|||||

Db 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKKNYCRNPD 60

QY 108 RRRPWCYVQVGLKPLVQECMVHDCAD 133
:|||||

Db 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 48

US-10-424-999-62

; Sequence 62, Application US/10424999

; Publication No. US20040052810A1

; GENERAL INFORMATION:

; APPLICANT: Nesbit, Mark

; APPLICANT: Cameron, Beatrice

; APPLICANT: Blanche, Francis

; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods fo

; FILE REFERENCE: ST01027-A

; CURRENT APPLICATION NUMBER: US/10/424,999

; CURRENT FILING DATE: 2003-04-29

; PRIOR FILING DATE: 2002-09-04

; NUMBER OF SEQ ID NOS: 70

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 62

; LENGTH: 86

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Human kringle domain ATF-Kringle (Abrogen)

US-10-424-999-62

Query Match 59.6%; Score 499; DB 12; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.4e-42;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKKNYCRNPD 107
:|||||
Db 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKKNYCRNPD 60

QY 108 RRRPWCYVQVGLKPLVQECMVHDCAD 133
:|||||

Db 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 49

US-10-425-000-97

; Sequence 97, Application US/10425000

; Publication No. US20040052777A1

; GENERAL INFORMATION:

; APPLICANT: Nesbit, Mark

; APPLICANT: Cameron, Beatrice

; APPLICANT: Blanche, Francis

; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit

; FILE REFERENCE: ST01027-B

; CURRENT APPLICATION NUMBER: US/10/425,000

; CURRENT FILING DATE: 2003-04-29

; PRIOR FILING DATE: 2002-09-04

; NUMBER OF SEQ ID NOS: 105

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 97

; LENGTH: 86

; TYPE: PRT

; OTHER INFORMATION: Human abrogen

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human kringle domain ATP-Kringle (Abrogen)
US-10-425-000-97

Query Match          59.6%; Score 499; DB 12; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.4e-42;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 KTCYEGNGHFYRGKASTDTMGRCLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPDN 107
Db 1 KTCYEGNGHFYRGKASTDTMGRCLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPDN 60

QY 108 RRRPWCYVQGLKPLVQECMWHDCAD 133
Db 61 RRRPWCYVQGLKPLVQECMWHDCAD 86

RESULT 50
US-10-233-675A-5
; Sequence 5, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
US-10-233-675A-5

Query Match          59.6%; Score 499; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.4e-42;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 KTCYEGNGHFYRGKASTDTMGRCLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPDN 107
Db 1 KTCYEGNGHFYRGKASTDTMGRCLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPDN 60

QY 108 RRRPWCYVQGLKPLVQECMWHDCAD 133
Db 61 RRRPWCYVQGLKPLVQECMWHDCAD 86

Search completed: May 25, 2004, 15:03:48
Job time : 27.0684 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:48:05 ; Search time 9.90096 Seconds
(without alignments)
745.636 Million cell updates/sec

Title: US-09-880-503-8

Perfect score: 837

Sequence: 1 SNEHQVPSNCDLNGTGV.....QECMVHDCADGKPSPEE 143

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.psp:*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.psp:*
- 3: /cgn2_6/ptodata/2/iaa/5A_COMB.psp:*
- 4: /cgn2_6/ptodata/2/iaa/5B_COMB.psp:*
- 5: /cgn2_6/ptodata/2/iaa/5A_COMB.psp:*
- 6: /cgn2_6/ptodata/2/iaa/5B_COMB.psp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	837	100.0	200	4 US-09-101-272G-73
2	837	100.0	208	4 US-09-101-272G-98
3	837	100.0	411	1 US-08-087-183-1
4	837	100.0	411	1 US-08-286-748B-18
5	837	100.0	411	1 US-08-153-799-18
6	837	100.0	411	4 US-09-403-736-2
7	837	100.0	430	1 US-07-942-157A-3
8	837	100.0	431	4 US-09-101-272G-1
9	837	100.0	431	6 518829-1
10	827	98.8	411	3 US-09-181-816-1
11	826	98.7	411	2 US-08-560-098A-48
12	821.5	98.1	430	6 5219569-2
13	813	97.1	157	3 US-08-142-590B-25
14	793	94.7	138	2 US-08-797-689-12
15	793	94.7	138	4 US-09-984-186-12
16	788	94.1	194	4 US-09-101-272G-80
17	788	94.1	201	4 US-09-101-272G-96
18	559	66.8	432	2 US-08-560-098A-47
19	558	66.7	365	1 US-08-093-741-83
20	558	66.7	365	1 US-08-720-012-83
21	558	66.7	393	2 US-08-560-098A-44
22	558	66.7	393	3 US-08-967-024C-24
23	558	66.7	393	3 US-08-967-024C-25
24	513	61.3	89	4 US-09-101-272G-62
25	335.5	40.1	477	2 US-08-560-098A-51
26	328.5	39.2	527	1 US-07-609-510B-16
27	328.5	39.2	527	2 US-08-811-949-39

RESULT 1
US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

ALIGNMENTS

28	328.5	39.2	527	5	PCT-US91-01025A-2	Sequence 2, Appl
29	328.5	39.2	527	6	5185259-8	Patent No. 5185259
30	328.5	39.2	527	6	5520913-1	Patent No. 5520913
31	328.5	39.2	546	6	5200340-6	Patent No. 5200340
32	328.5	39.2	562	2	US-08-811-949-43	Sequence 43, Appl
33	328.5	39.2	562	2	US-08-560-098A-50	Sequence 50, Appl
34	328.5	39.2	562	2	US-08-883-795A-38	Sequence 38, Appl
35	328.5	39.2	562	4	US-09-703-695A-4	Sequence 4, Appl
36	328.5	39.2	562	6	5185259-3	Patent No. 5185259
37	328.5	39.2	562	6	5200340-2	Patent No. 5200340
38	328.5	39.2	562	6	5344773-2	Patent No. 5344773
39	321.5	38.4	562	6	5244676-5	Patent No. 5244676
40	300	35.8	233	3	US-08-438-745-15	Sequence 15, Appl
41	300	35.8	233	3	US-08-438-745-17	Sequence 17, Appl
42	300	35.8	233	3	US-09-219-019-15	Sequence 15, Appl
43	300	35.8	233	3	US-09-219-019-17	Sequence 17, Appl
44	300	35.8	233	5	PCT-US94-05669A-15	Sequence 15, Appl
45	300	35.8	233	5	PCT-US94-05669A-17	Sequence 17, Appl
46	300	35.8	235	3	US-08-438-745-13	Sequence 13, Appl
47	300	35.8	235	3	US-09-219-019-13	Sequence 13, Appl
48	300	35.8	235	5	PCT-US94-05669A-13	Sequence 13, Appl
49	293	35.0	49	4	US-08-747-915-5	Sequence 5, Appl
50	293	35.0	49	4	US-09-285-783-5	Sequence 5, Appl
51	288	34.4	48	3	US-09-219-019-22	Sequence 22, Appl
52	268	32.0	44	2	US-08-747-915-1	Sequence 1, Appl
53	268	32.0	44	4	US-09-285-783-1	Sequence 1, Appl
54	262	31.3	655	1	US-08-148-910-12	Sequence 12, Appl
55	262	31.3	655	1	US-08-448-937A-12	Sequence 12, Appl
56	230.5	27.5	437	2	US-08-811-949-49	Sequence 49, Appl
57	230.5	27.5	437	2	US-08-811-949-51	Sequence 51, Appl
58	230.5	27.5	437	2	US-08-811-949-55	Sequence 55, Appl
59	230.5	27.5	437	2	US-08-811-949-57	Sequence 57, Appl
60	230.5	27.5	472	2	US-08-811-949-63	Sequence 63, Appl
61	224.5	26.8	356	1	US-08-427-640-8	Sequence 8, Appl
62	234	26.8	326	4	US-09-411-977-3	Sequence 3, Appl
63	220.5	26.3	378	4	US-09-553-498-10	Sequence 10, Appl
64	220.5	26.3	378	4	US-09-618-869-10	Sequence 10, Appl
65	219.5	26.2	389	2	US-08-811-949-67	Sequence 67, Appl
66	218.5	26.1	355	1	US-08-137-116-1	Sequence 1, Appl
67	218.5	26.1	355	1	US-08-217-618-1	Sequence 1, Appl
68	218.5	26.1	355	1	US-08-427-640-2	Sequence 2, Appl
69	218.5	26.1	355	1	US-08-217-617A-1	Sequence 1, Appl
70	218.5	26.1	355	1	US-08-217-616-1	Sequence 1, Appl
71	218.5	26.1	355	2	US-08-811-949-45	Sequence 45, Appl
72	218.5	26.1	355	2	US-08-811-949-47	Sequence 47, Appl
73	218.5	26.1	355	2	US-08-811-949-53	Sequence 53, Appl
74	218.5	26.1	355	3	US-08-794-528-1	Sequence 1, Appl
75	218.5	26.1	355	6	5223256-1	Patent No. 5223256

```
; OTHER INFORMATION: ATF domain of uPA
US-09-101-272G-73

Query Match      100.0%; Score 837; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 6.1e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCHEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHVPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCHEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKKNYCRNPNRRRPPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKKNYCRNPNRRRPPWCYVQVGLK 140
QY 121 PLVQECMWHDCADGKKPSPPEE 143
DB 141 PLVQECMWHDCADGKKPSPPEE 163

RESULT 2
US-09-101-272G-98
; Sequence 98, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 98
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFI-ML chimeric protein
US-09-101-272G-98

Query Match      100.0%; Score 837; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 6.4e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCHEIDKSKTCYEGNGHFYRG 60
DB 2 SNELHVPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCHEIDKSKTCYEGNGHFYRG 61
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKKNYCRNPNRRRPPWCYVQVGLK 120
DB 62 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKKNYCRNPNRRRPPWCYVQVGLK 121
QY 121 PLVQECMWHDCADGKKPSPPEE 143
DB 122 PLVQECMWHDCADGKKPSPPEE 144

RESULT 3
US-08-087-163-1
; Sequence 1, Application US/08087163
; Patent No. 5472692
; GENERAL INFORMATION:
; APPLICANT: Liu, Jian-Ning
; APPLICANT: Gurewicz, Victor
; TITLE OF INVENTION: PRO-UKINASE MUTANTS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
```

```
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,163
; FILING DATE: 07/02/93
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04353/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
US-08-087-163-1

Query Match      100.0%; Score 837; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.4e-75;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCHEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHVPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCHEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKKNYCRNPNRRRPPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKKNYCRNPNRRRPPWCYVQVGLK 120
QY 121 PLVQECMWHDCADGKKPSPPEE 143
DB 121 PLVQECMWHDCADGKKPSPPEE 143

RESULT 4
US-08-286-748B-18
; Sequence 18, Application US/08286748B
; Patent No. 5755542
; GENERAL INFORMATION:
; APPLICANT: Victor Gurewicz
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
; OF DRUGS BY PLATELETS FOR THE TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,748B
; FILING DATE: August 5, 1994
; CLASSIFICATION: 424
```

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: J. Peter Fasse
;; REGISTRATION NUMBER: 32,983
;; REFERENCE/DOCKET NUMBER: 04547/013001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 542-5070
;; TELEFAX: (617) 542-8906
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 18:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 411
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-286-748B-18

Query Match 100.0%; Score 837; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.4e-75;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
DB 61 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPKSPPEE 143
DB 121 PLVQECMVHDCADGKPKSPPEE 143

RESULT 5
US-08-153-799-18
; Sequence 18, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; APPLICATION INFORMATION:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Swope, R Hain
;; REGISTRATION NUMBER: 24964
;; REFERENCE/DOCKET NUMBER: 92H832
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (908) 665 2400
;; TELEFAX: (908) 771 6159
;; TELEX: 219484
;; INFORMATION FOR SEQ ID NO: 18:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 411 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-153-799-18

Query Match 100.0%; Score 837; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.4e-75;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
DB 61 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPKSPPEE 143
DB 121 PLVQECMVHDCADGKPKSPPEE 143

RESULT 6
US-09-403-736-2
; Sequence 2, Application US/09403736
; Patent No. 6638502
; GENERAL INFORMATION:
; APPLICANT: Aventis S.A.
; APPLICANT: Li, Hong
; APPLICANT: LU, He
; APPLICANT: GRISCELLI, Frank
; APPLICANT: OPOLO, Paule
; APPLICANT: SORIA, Claudine
; APPLICANT: RAGOT, Thierry
; APPLICANT: LEGRAND, Yves
; APPLICANT: SORIA, Jeanette
; APPLICANT: MABILAT, Christelle
; APPLICANT: PERRICAUDET, Michel
; APPLICANT: YEH, Patrice
; TITLE OF INVENTION: Adenovirus-Mediated Intratumoral Delivery Of An Angiogenesis Ant
; FILE REFERENCE: A2778A-US
; CURRENT APPLICATION NUMBER: US/09/403,736
; CURRENT FILING DATE: 1999-10-26
; PRIOR APPLICATION NUMBER: PCT/EP98/02491
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/044,980
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: humanurokinase
US-09-403-736-2

Query Match 100.0%; Score 837; DB 4; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.4e-75;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60

Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRCLPWSATVLQOYTHAHRSDALQGLGKNCNPNRRPWCYVQVGLK 120
Db 61 KASTDTMGRCLPWSATVLQOYTHAHRSDALQGLGKNCNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPKSPPEE 143
Db 121 PLVQECMVHDCADGKPKSPPEE 143

RESULT 7

US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,157A
; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631673
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TS1108Cont.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)815-6508
; TELEFAX: (404)815-6555
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..19
; OTHER INFORMATION: /label= peptide
; OTHER INFORMATION: /note= "MAP signal"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 198..203
; OTHER INFORMATION: /label= modified
; OTHER INFORMATION: /note= "six amino acids deleted in mutant"
US-07-942-157A-3

Query Match 100.0%; Score 837; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.5e-75;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 20 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 79
QY 61 KASTDTMGRCLPWSATVLQOYTHAHRSDALQGLGKNCNPNRRPWCYVQVGLK 120

Db 80 KASTDTMGRCLPWSATVLQOYTHAHRSDALQGLGKNCNPNRRPWCYVQVGLK 139
QY 121 PLVQECMVHDCADGKPKSPPEE 143
Db 140 PLVQECMVHDCADGKPKSPPEE 162

RESULT 8

US-09-101-272G-1
; Sequence 1, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissan Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (21)..()
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (20)..()
; OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)
US-09-101-272G-1

Query Match 100.0%; Score 837; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.5e-75;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRCLPWSATVLQOYTHAHRSDALQGLGKNCNPNRRPWCYVQVGLK 120
Db 81 KASTDTMGRCLPWSATVLQOYTHAHRSDALQGLGKNCNPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKPKSPPEE 143
Db 141 PLVQECMVHDCADGKPKSPPEE 163

RESULT 9

5188829-1
; Patent No. 5188829
; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEXI; YAMADA, CHIKAKO
; TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1988
; SEQ ID NO:1:
; LENGTH: 431
5188829-1

Query Match 100.0%; Score 837; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.5e-75;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPPWCYVOVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPPWCYVOVGLK 140
QY 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 141 PLVQECMVHDCADGKKPSSPPEE 163

RESULT 10

US-09-181-816-1
; Sequence 1, Application US/09181816
; Patent No. 6277818
; GENERAL INFORMATION:
; APPLICANT: MAZAR, Andrew P.
; APPLICANT: JONES, Terence R.
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
; TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR
; FILE REFERENCE: 329042000300 SIDN 1-7
; CURRENT APPLICATION NUMBER: US/09/181,816
; CURRENT FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 411
; TYPE: PRI
; ORGANISM: Homo sapiens
US-09-181-816-1

Query Match 98.8%; Score 827; DB 3; Length 411;
Best Local Similarity 99.3%; Pred. No. 1.4e-74;
Matches 142; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCHEIDKSKTCYEGNGHFYRG 60
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCHEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPPWCYVOVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPPWCYVOVGLK 120
QY 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 121 PLVQECMVHDCADGKKPSSPPEE 143

RESULT 11

US-08-560-098A-48
; Sequence 48, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEBTD, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-48

Query Match 98.7%; Score 826; DB 2; Length 411;
Best Local Similarity 99.3%; Pred. No. 1.7e-74;
Matches 142; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCHEIDKSKTCYEGNGHFYRG 60
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCHEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPPWCYVOVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPPWCYVOVGLK 120
QY 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 121 PLVQECMVHDCADGKKPSSPPEE 143

RESULT 12

5219569-2
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO: 2
; LENGTH: 430
5219569-2

Query Match 98.1%; Score 821.5; DB 6; Length 430;
Best Local Similarity 99.3%; Pred. No. 5.1e-74;
Matches 142; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCHEIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCHEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPPWCYVOVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDNRPPWCYVOVGLK 140
QY 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 141 PLVQECMVHDCADG-KPSSPPEE 162

RESULT 13

US-08-142-5908-25
; Sequence 25, Application US/081425908
; Patent No. 6120765

GENERAL INFORMATION:
APPLICANT: HIBINO, Toshihiko; TAKAHASHI, Tadahito; HORII, Izumi; and GOSTINCK,
TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,590B
FILING DATE: 25-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,318
FILING DATE: 02-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-009CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-142-590B-25

Query Match 97.1%; Score 813; DB 3; Length 157;
Best Local Similarity 97.2%; Pred. No. 1.1e-73;
Matches 139; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGSPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
DB 61 KASTDTMGSPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPKSPPEE 143
DB 121 PLVQECMVHDCADGKPKSPPEE 143

RESULT 14
US-08-797-689-12
Sequence 12, Application US/08797689
Patent No. 5876969
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville

GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville

STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-797-689-12

Query Match 94.7%; Score 793; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 9.7e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60
DB 4 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGSPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
DB 64 KASTDTMGSPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPDNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 15
US-09-984-186-12
Sequence 12, Application US/09984186
Patent No. 6686179
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426


```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-720-012-83

Query Match          66.7%; Score 558; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 7.8e-48;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPD 106
Db 1 SKTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPD 60

QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 143
Db 61 NRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 97

RESULT 21
US-08-560-098A-44
; Sequence 44, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNENDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560.098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-44

Query Match          66.7%; Score 558; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 8.5e-48;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPD 106
Db 2 SKTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPD 61

QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 143
```

```
Db 62 NRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 98

RESULT 22
US-08-967-024C-24
; Sequence 24, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNENDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967.024C
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 42 665.8
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-967-024C-24

Query Match          66.7%; Score 558; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 8.5e-48;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPD 106
Db 2 SKTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPD 61

QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 143
Db 62 NRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 98

RESULT 23
US-08-967-024C-25
; Sequence 25, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNENDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
```

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
;; STREET: 1200 G Street, N.W., Suite 700
;; CITY: Washington
;; STATE: DC
;; ZIP: 20005
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/967,024C
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA: P 44 42 665.8
;; FILING DATE: 30-NOV-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: EVANS, Joseph D.
;; REGISTRATION NUMBER: 26,269
;; REFERENCE/DOCKET NUMBER: 148/42444
;; TELEPHONE: (202) 628-8800
;; TELEFAX: (202) 628-8844
;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 393 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-967-024C-25

Query Match 66.7%; Score 558; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 8.5e-48;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNVC 106
DB 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNVC 61

QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPPE 143
DB 62 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPPE 98

RESULT 24
US-09-101-272G-62
; Sequence 62, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 62
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: residues 43-131 of the ATF domain of uPA
US-09-101-272G-62

Query Match 61.3%; Score 513; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 4.8e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 61.3%; Score 513; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 4.8e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 EIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNVC 102
DB 1 EIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNVC 60

QY 103 RNPDRRPPWCYVQVGLKPLVQECMVHDC 131
DB 61 RNPDRRPPWCYVQVGLKPLVQECMVHDC 89

RESULT 25
US-08-560-098A-51
; Sequence 51, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-51

Query Match 40.1%; Score 335.5; DB 2; Length 477;
Best Local Similarity 46.3%; Pred. No. 1.6e-25;
Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;

QY 3 ELHVP-----SNCDLNGGTCVKNKYFNSIHWNCNPKFGQHCIEDKSKTCYEGNGHFY 58
DB 78 QCHTVPVKSCSELRCFNGGTCWQASPSDF-VQCQPKGYTKQCEVDTHATCYKDGQVY 136

QY 59 RGKASTDTMGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNVCNPNDRRPPWCYVQV 118
DB 137 RGTWSTESGAQCINMNSNLLTRTYNGRSDAITLQGNHNYCRPNDRRPPWCYVQV 196

QY 119 LKPLVQECMVHDC 132
DB 197 SKFILEPCSPVCS 210

RESULT 26

Query Match 40.1%; Score 335.5; DB 2; Length 477;
Best Local Similarity 46.3%; Pred. No. 1.6e-25;
Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-811-949-39

Query Match 39.2%; Score 328.5; DB 2; Length 527;
Best Local Similarity 46.3%; Pred. No. 8.9e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHVP-SNCD---LNGGTCSVKNYFSNIHWNCNCPKFGGCHCEIDKSKTCYEGNGHFY 58
DB 42 QCHSVPVYKSCSEPRCNGTTCQALYFSDF-VCCPEGFAGKCCSIDTRATCYEDQGGSY 100
QY 59 RGKASTDMGRCLPWN SATVLQQTVAHRS DALQLGLGKKNYCNPNRRRPMCYVQVG 118
DB 101 RGTWSTAESGAECTNNSSALAKPYSGRRPDAIRLGLGNHNYCNRPDRD SKPCYVFK 160
QY 119 LKPLVQECMVHDCADG 134
DB 161 GKYSEFCSTPACSEG 176

RESULT 28
PCT-US91-01025A-2
; Sequence 2, Application PC/TUS9101025A
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
; TITLE OF INVENTION: Specific Properties
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01025A
; FILING DATE: 19910214
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/486,657
; FILING DATE: 1 March 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 454P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/366-1896
; TELEFAX: 415/952-9881
; TELEFAX: 910/371-7168

```

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US91-01025A-2

Query Match 39.2%; Score 328.5; DB 5; Length 527;
Best Local Similarity 46.3%; Pred. No. 8.9e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHOVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHGY 58
DB 42 QCHSVPVKSCSEPCFCFNGGTCQALYFSDP-VQCPEGFAGKCCCEIDTRATCYEDQGISY 100

QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRERRPWCYVQVG 118
DB 101 RGTWSTAESGAECTNWNSSALAAQPYSGRRPDALRLGLGNHNYCRNPDNRDRSKPWCYVFXA 160

QY 119 LKPLVQECMVHDCADG 134
DB 161 GKYSEFCSTPACSEG 176

RESULT 29
5185259-8
Patent No. 5185259
APPLICANT: GOEDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
VEHAR, GORDON A.
TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
ACTIVATOR
NUMBER OF SEQUENCES: 15
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/489,855
FILING DATE: 02-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 12,694
FILING DATE: 09-FEB-1987
APPLICATION NUMBER: 483,052
FILING DATE: 07-APR-1983
APPLICATION NUMBER: 398,003
FILING DATE: 14-JUL-1982
APPLICATION NUMBER: 374,860
FILING DATE: 05-MAY-1982
SEQ ID NO: 8
LENGTH: 527

Query Match 39.2%; Score 328.5; DB 6; Length 527;
Best Local Similarity 46.3%; Pred. No. 8.9e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHOVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHGY 58
DB 42 QCHSVPVKSCSEPCFCFNGGTCQALYFSDP-VQCPEGFAGKCCCEIDTRATCYEDQGISY 100

QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRERRPWCYVQVG 118
DB 101 RGTWSTAESGAECTNWNSSALAAQPYSGRRPDALRLGLGNHNYCRNPDNRDRSKPWCYVFXA 160

QY 119 LKPLVQECMVHDCADG 134
DB 161 GKYSEFCSTPACSEG 176

RESULT 30
5185259-8
Patent No. 5185259
APPLICANT: ANDERSON, STEPHEN; BENNETT, WILLIAM F.; BORTSTEIN,
DAVID; HIGGINS, DEBORAH L.; PAONI, NICHOLAS F.; ZOLLER, MARK J.
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
ZYMOTIC PROPERTIES
NUMBER OF SEQUENCES: 35

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/88,451
FILING DATE: 06-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 770,510
FILING DATE: 03-OCT-1991
APPLICATION NUMBER: 384,608
FILING DATE: 24-JUL-1989
APPLICATION NUMBER: 240,856
FILING DATE: 02-SEP-1986
SEQ ID NO: 1
LENGTH: 527
5520913-1

Query Match 39.2%; Score 328.5; DB 6; Length 527;
Best Local Similarity 46.3%; Pred. No. 8.9e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHOVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHGY 58
DB 42 QCHSVPVKSCSEPCFCFNGGTCQALYFSDP-VQCPEGFAGKCCCEIDTRATCYEDQGISY 100

QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRERRPWCYVQVG 118
DB 101 RGTWSTAESGAECTNWNSSALAAQPYSGRRPDALRLGLGNHNYCRNPDNRDRSKPWCYVFXA 160

QY 119 LKPLVQECMVHDCADG 134
DB 161 GKYSEFCSTPACSEG 176

RESULT 31
5200340-6
Patent No. 5200340
APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,
PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI
TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
ACTIVATORS
NUMBER OF SEQUENCES: 34
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/53,412
FILING DATE: 22-MAY-1987
SEQ ID NO: 6
LENGTH: 546
5200340-5

Query Match 39.2%; Score 328.5; DB 6; Length 546;
Best Local Similarity 46.3%; Pred. No. 9.2e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHOVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHGY 58
DB 77 QCHSVPVKSCSEPCFCFNGGTCQALYFSDP-VQCPEGFAGKCCCEIDTRATCYEDQGISY 135

QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRERRPWCYVQVG 118
DB 136 RGTWSTAESGAECTNWNSSALAAQPYSGRRPDALRLGLGNHNYCRNPDNRDRSKPWCYVFXA 195

QY 119 LKPLVQECMVHDCADG 134
DB 196 GKYSEFCSTPACSEG 211

RESULT 32
US-08-811-949-43
Sequence 43, Application US/08811949
Patent No 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI

APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESSES:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM: disk

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/811,949

FILING DATE: 05-MAR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 18-966-0

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

LENGTH: 562 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-811-949-43

Query Match 39.2%; Score 328.5; DB 2; Length 562;

Best Local Similarity 46.3%; Pred. No. 9.5e-25;

Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHOVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHVF 58

Db 77 QCHSVPVKSGSEPRCFNGGTCQALYFSDP-VQCPEGFAGKCCETDTRATCYEQGISY 135

QY 59 RGKASTDTMGRCLPWNSTVLTQOYHAHRSALQLGLGKHNYCRNPNRRPWCYQVVG 118

Db 136 RGTWSTAESGAECTNWNSSALAKPYSGRRPDALRLGLGNHNYCRNPNDRDSKPCWYVFK 195

QY 119 LKPLVQECMVHDCADG 134

Db 196 GKYSSEFCSTPACSEG 211

RESULT 33

US-08-560-098A-50

Sequence 50, Application US/08560098A

Patent No. 5976841

GENERAL INFORMATION:

APPLICANT: WENNDT, Stephan

APPLICANT: HEINZEL-WIELAND, Regina

APPLICANT: STEFFENS, Gerd Josef

TITLE OF INVENTION: Proteins having Fibrinolytic and

NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan

STREET: 1200 G Street, N.W., Suite 700

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/560,098A

FILING DATE: 17-NOV-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: P 44 40 892.7

FILING DATE: 17-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 148/42448

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800

TELEFAX: (202) 628-8844

INFORMATION FOR SEQ ID NO: 50:

SEQUENCE CHARACTERISTICS:

LENGTH: 562 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-560-098A-50

Query Match 39.2%; Score 328.5; DB 2; Length 562;

Best Local Similarity 46.3%; Pred. No. 9.5e-25;

Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHOVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHVF 58

Db 77 QCHSVPVKSGSEPRCFNGGTCQALYFSDP-VQCPEGFAGKCCETDTRATCYEQGISY 135

QY 59 RGKASTDTMGRCLPWNSTVLTQOYHAHRSALQLGLGKHNYCRNPNRRPWCYQVVG 118

Db 136 RGTWSTAESGAECTNWNSSALAKPYSGRRPDALRLGLGNHNYCRNPNDRDSKPCWYVFK 195

QY 119 LKPLVQECMVHDCADG 134

Db 196 GKYSSEFCSTPACSEG 211

RESULT 34

US-08-883-795A-38

Sequence 38, Application US/08883795A

Patent No. 5985607

GENERAL INFORMATION:

APPLICANT: Delcuve, Genevieve

APPLICANT: Awang, Gregor

TITLE OF INVENTION: Recombinant DNA Molecules and Expression

TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: BERESKIN & PARR

STREET: 40 King Street West

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5H 3Y2

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/883,795A

FILING DATE: 27-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Gravelle, Micheline

REGISTRATION NUMBER: 40,261

REFERENCE/DOCKET NUMBER: 7841-062

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 364-7311


```
; APPLICATION NUMBER: US/07/782,686
; FILING DATE: 01-OCT-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 656,770
; FILING DATE: 01-OCT-1984
; SEQ ID NO:2:
; LENGTH: 562
; 5344773-2

Query Match      39.2%; Score 328.5; DB 6; Length 562;
Best Local Similarity 46.3%; Pred. No. 9.5e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQP-SNCD----CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYF 58
DB 77 QCHSVPVKSCSEPCFNGGTCQQAIFYSDP-VQCPEGFAGKCEIDTATCYEQGISY 135
QY 59 RGKASTDTMRPCLPWNASVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVG 118
DB 136 RGTWSTAESAECTWNSSALAKPYSGRPDAIRLGLGNHNYCRNPDRDRSKFWCVFKA 195
QY 119 LKPLVQECMVHDCADG 134
DB 196 GKYSSEFCSTPACSEG 211

RESULT 39
5244676-5
; Patent No. 5244676
; APPLICANT: BELL, LESLIE D.; MAYER, ERNEST J.; PALMIER, MARK O.
; TOLUNAY, H.ESER; WARREN, THOMAS G.; WUN, TZE-CHIN
; TITLE OF INVENTION: MODIFIED TISSUE PLASMINOGEN ACTIVATOR
; WITH MODIFIED GLYCOSYLATION SITE
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/203,047
; FILING DATE: 06-JUN-1988
; SEQ ID NO:5:
; LENGTH: 562
; 5244676-5

Query Match      38.4%; Score 321.5; DB 6; Length 562;
Best Local Similarity 44.9%; Pred. No. 4.8e-24;
Matches 61; Conservative 16; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYF 58
DB 77 QCHSVPVKSCSEPCFNGGTCQQAIFYSDP-VQCPEGFAGKCEIDGNSDCYFGSGAY 135
QY 59 RGKASTDTMRPCLPWNASVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVG 118
DB 136 RGTWSTAESAECTPWNSSMILIGKVTYQNPSPAQALGLGKHNYCRNPDPGDAKFWCHVLKN 195
QY 119 LKPLVQECMVHDCADG 134
DB 196 RLITWEYCDVPSCSEG 211

RESULT 40
US-08-438-745-15
; Sequence 15, Application US/08438745
; Patent No. 6248715
; GENERAL INFORMATION:
; APPLICANT: Rosenberg, Steven
; APPLICANT: Stratton-Thomas, Jennifer
; TITLE OF INVENTION: Expression of Urokinase Plasminogen
; TITLE OF INVENTION: Activator Inhibitors
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,745
; FILING DATE: 10-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/070,153
; FILING DATE: 01-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
```

```
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,745
; FILING DATE: 10-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/070,153
; FILING DATE: 01-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0939.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-438-745-15

Query Match      35.8%; Score 300; DB 3; Length 233;
Best Local Similarity 77.9%; Pred. No. 2.4e-22;
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 1 SNELHQPVSNCDCNLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYF 60
DB 29 SNELHQPVSNCDCNLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKQGGSGDFDYE 88
QY 61 KASDTMG 68
DB 89 KMANANKG 96

RESULT 41
US-08-438-745-17
; Sequence 17, Application US/08438745
; Patent No. 6248715
; GENERAL INFORMATION:
; APPLICANT: Rosenberg, Steven
; APPLICANT: Stratton-Thomas, Jennifer
; TITLE OF INVENTION: Expression of Urokinase Plasminogen
; TITLE OF INVENTION: Activator Inhibitors
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,745
; FILING DATE: 10-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/070,153
; FILING DATE: 01-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
```

REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0939.001
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-438-745-17

Query Match 35.8%; Score 300; DB 3; Length 233;
Best Local Similarity 77.9%; Pred. No. 2.4e-22;
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKPGGQHCIDKSKTCYEGNGHFYRG 60
DB 29 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKPGGQHCIDKSKTCYEGNGHFYRG 88
QY 61 KASTDTMG 68
DB 89 KWANANKG 96

RESULT 42
US-09-219-019-15
Sequence 15, Application US/09219019
Patent No. 6268341
GENERAL INFORMATION:
APPLICANT: ROSENBERG, STEVEN
TITLE OF INVENTION: EXPRESSION OF UROKINASE PLASMINOGEN ACTIVATOR
TITLE OF INVENTION: INHIBITORS
FILE REFERENCE: 23533-0005
CURRENT APPLICATION NUMBER: US/09/219,019
CURRENT FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 08/438,263
PRIOR FILING DATE: 1995-05-10
PRIOR APPLICATION NUMBER: 08/280,288
PRIOR FILING DATE: 1994-07-26
PRIOR APPLICATION NUMBER: 08/070,153
PRIOR FILING DATE: 1993-06-01
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 233
TYPE: PRT
ORGANISM: Homo sapiens
US-09-219-019-15

Query Match 35.8%; Score 300; DB 3; Length 233;
Best Local Similarity 77.9%; Pred. No. 2.4e-22;
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKPGGQHCIDKSKTCYEGNGHFYRG 60
DB 29 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKPGGQHCIDKSKTCYEGNGHFYRG 88
QY 61 KASTDTMG 68
DB 89 KWANANKG 96

RESULT 43
US-09-219-019-17
Sequence 17, Application US/09219019
Patent No. 6268341
GENERAL INFORMATION:
APPLICANT: ROSENBERG, STEVEN
TITLE OF INVENTION: EXPRESSION OF UROKINASE PLASMINOGEN ACTIVATOR

TITLE OF INVENTION: INHIBITORS
FILE REFERENCE: 23533-0005
CURRENT APPLICATION NUMBER: US/09/219,019
CURRENT FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 08/438,263
PRIOR FILING DATE: 1995-05-10
PRIOR APPLICATION NUMBER: 08/280,288
PRIOR FILING DATE: 1994-07-26
PRIOR APPLICATION NUMBER: 08/070,153
PRIOR FILING DATE: 1993-06-01
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 17
LENGTH: 233
TYPE: PRT
ORGANISM: Homo sapiens
US-09-219-019-17

Query Match 35.8%; Score 300; DB 3; Length 233;
Best Local Similarity 77.9%; Pred. No. 2.4e-22;
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKPGGQHCIDKSKTCYEGNGHFYRG 60
DB 29 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKPGGQHCIDKSKTCYEGNGHFYRG 88
QY 61 KASTDTMG 68
DB 89 KWANANKG 96

RESULT 44
PCT-US94-05669A-15
Sequence 15, Application PC/TUS9405669A
GENERAL INFORMATION:
APPLICANT: Chiron Corporation
TITLE OF INVENTION: Expression of Urokinase Plasminogen
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05669A
FILING DATE: 19-MAY-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0939.100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-05669A-15

Query Match 35.8%; Score 300; DB 5; Length 233;
Best Local Similarity 77.9%; Pred. No. 2.4e-22;
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

Best Local Similarity 77.9%; Pred. No. 2.5e-22;
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 31 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 90
QY 61 KASTDTMG 68
Db 91 KMANANKG 98
RESULT 48
PCT-US94-05669A-13
US-08-747-915-5
; Sequence 13, Application PC/TUS9405669A
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Expression of Urokinase Plasminogen
; Activator Inhibitors
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05669A
; FILING DATE: 19-MAY-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0939.100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-855-3542
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-05669A-13
Query Match 35.8%; Score 300; DB 5; Length 235;
Best Local Similarity 77.9%; Pred. No. 2.5e-22;
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 31 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 90
QY 61 KASTDTMG 68
Db 91 KMANANKG 98
RESULT 49
US-08-747-915-5
; Sequence 5, Application US/08747915
; Patent No. 5942492
; GENERAL INFORMATION:
; APPLICANT: Jones, Terence R.
; APPLICANT: Haney, David N.
; APPLICANT: Varga, Janos
; TITLE OF INVENTION: CYCLIC PEPTIDES THAT BIND TO

; TITLE OF INVENTION: UROKINASE-TYPE PLASMINOGEN ACTIVATOR RECEPTOR
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVENUE, NW
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,915
; FILING DATE: 12-NOV-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 32904-20001.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030 MRSNFOERSWSH
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Disulfide-bond
; LOCATION: group(11..19, 13..31, 33..42)
US-08-747-915-5
Query Match 35.0%; Score 293; DB 2; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.1e-22;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 49
Db 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 49
RESULT 50
US-09-285-783-5
; Sequence 5, Application US/09285783
; Patent No. 6544710
; GENERAL INFORMATION:
; APPLICANT: Jones, Terence R.
; APPLICANT: Haney, David N.
; APPLICANT: Varga, Janos
; TITLE OF INVENTION: CYCLIC PEPTIDES THAT BIND TO
; UROKINASE-TYPE PLASMINOGEN ACTIVATOR RECEPTOR
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RADER, FISMAN & GRAUER
; STREET: 1233 20TH STREET NW, SUITE 501
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/285,783
; FILING DATE: 05-Apr-1999
; CLASSIFICATION: <Unknown>

```

; ATTORNEY/AGENT INFORMATION:
; NAME: LIVNAT, SHMUEL
; REGISTRATION NUMBER: 33,949
; REFERENCE/DOCKET NUMBER: ANG-001/DIV (80144-0007)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 955-8787
; TELEFAX: (202) 955-3751
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Disulfide-bond
; LOCATION: group(11..19, 13..31, 33..42)
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-285-783-5

Query Match      35.0%; Score 293; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.1e-22;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKPKFGGQCHCEIDKSKT 49
Db      1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKPKFGGQCHCEIDKSKT 49

Search completed: May 25, 2004, 15:00:08
Job time : 11.901.secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:47:10 ; Search time 5.55302 Seconds
(without alignments)
1662.947 Million cell updates/sec

Title: US-09-880-503-9

Perfect score: 554

Sequence: 1 KTCYEGNHFRGKASTDTM.....QECVHDCADGKKPSPPEE 96

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

PIR 78.*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	554	100.0	431	1 UKHU	u-plasminogen acti
2	519	93.7	433	1 UKBAY	u-plasminogen acti
3	437.5	79.0	442	1 UKPG	u-plasminogen acti
4	427	77.1	432	1 SI8932	u-plasminogen acti
5	422	76.2	433	1 JN0360	u-plasminogen acti
6	408	73.6	433	1 UKMS	u-plasminogen acti
7	241	43.5	431	2 JS0599	t-plasminogen acti
8	241	43.5	477	1 A34369	t-plasminogen acti
9	241	43.5	477	2 JS0598	t-plasminogen acti
10	228.5	41.2	559	1 A35029	t-plasminogen acti
11	226	40.8	291	2 I38098	t-plasminogen acti
12	226	40.8	562	1 UKHUT	t-plasminogen acti
13	220	39.7	394	2 JS0600	t-plasminogen acti
14	219	39.5	559	1 A29941	t-plasminogen acti
15	213	38.4	477	2 JS0597	t-plasminogen acti
16	210	37.9	558	2 JC5878	plasma hyaluronan-
17	199.5	36.0	560	1 JC4795	plasma hyaluronan-
18	198.5	35.8	434	1 A35005	hepatocyte growth
19	194.5	35.1	655	1 A46588	coagulation factor
20	191.5	34.6	603	2 S28941	coagulation factor
21	170.5	30.8	615	1 KFHU12	apolipoprotein(a)
22	161	29.1	1420	2 A32869	apolipoprotein(a)
23	157	28.3	4548	1 S00657	plasmin (EC 3.4.21
24	156	28.2	120	2 E61545	coagulation factor
25	155.5	28.1	593	2 A45201	plasmin (EC 3.4.21
26	153	27.6	89	2 A60140	plasmin (EC 3.4.21
27	153	27.6	460	2 E61545	plasmin (EC 3.4.21
28	152	27.4	812	1 PLBO	plasmin (EC 3.4.21
29	150	27.1	123	2 C61545	plasmin (EC 3.4.21

ALIGNMENTS

RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human

N:Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminoge
N:Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen ac
in form

C:Species: Homo sapiens (man)

C>Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000

C:Accession: A00931; I52209; J0102; A37561; I38102; S65783; A37562; A37563; A37564; A3

R:Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasi, F.

Nucleic Acids Res. 13, 2759-2771, 1985

A:Title: The human urokinase-plasminogen activator gene and its promoter.

A:Reference number: A00931; PMID:85215647; PMID:2987867

A:Accession: A00931

A:Molecule type: DNA

A:Residues: 1-431 <RIC>

A:Cross-references: GB:X02419; NID:G37601; PIDN:CAA26268.1; PID:q1834524

A:Note: the authors translated the codon ATG for residue 214 as Ile

R:Nagamine, Y.; Pearson, D.; Grattan, M.

Biochem. Biophys. Res. Commun. 132, 563-569, 1985

A:Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcine

A:Reference number: I52209; MUID:86050639; PMID:3933505

plasmin (EC 3.4.21
plasmin (EC 3.4.21
macrophage-stimula
plasmin (EC 3.4.21
plasmin (EC 3.4.21
macrophage-stimula
neurotrophic recep
plasmin (EC 3.4.21
neurotrophic recep
hypothetical prote
plasmin (EC 3.4.21
apolipoprotein(a)
plasmin (EC 3.4.21
neurotrophic recep
macrophage-stimula
hepatocyte growth
hepatocyte growth
hepatocyte growth
hepatocyte growth
plasmin precursor
hepatocyte growth
thrombin (EC 3.4.2
thrombin (EC 3.4.2
thrombin (EC 3.4.2
thrombin (EC 3.4.2
brain-specific ser
hypothetical prote
2,4-dichlorophenol
extracellular matr
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
anther-specific pr
hypothetical prote
hypothetical prote
proline-rich pepti
hypothetical prote
Niemann-Pick C dis
rnsd protein precu
apolipoprotein N-a
hypothetical prote
two-component sens
probable apolipop
homeotic protein H

A;Accession: I52209
A;Status: preliminary; translated from GB/EMBL/DBD
A;Molecule type: DNA
A;Residues: 145-161 <NAG1>
A;Cross-references: GB:X03027; NID:G340174; PIDN:AAA61257.1; PID:G340175
R;Nagai, M.; Hiramatsu, K.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama, Gene 36, 183-188, 1985
A;Title: Molecular cloning of cDNA coding for human preprourokinase.
A;Reference number: J70102; MUID:86056954; PMID:2415429
A;Accession: J70102
A;Molecule type: mRNA
A;Residues: 1-213, 'I', 215-431 <NAG2>
A;Cross-references: GB:X03226; NID:G340155; PIDN:AACT97138.1; PID:G340158; GB:D00244; NID:R;Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Biasi, F. Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A;Title: Identification and primary sequence of an unspliced human urokinase poly(A) + RN
A;Reference number: A37561; MUID:84272706; PMID:6589620
A;Accession: A37561
A;Molecule type: mRNA
A;Residues: 66-431 <VER>
A;Cross-references: GB:D00244; NID:G220138
R;Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elsen, DNA 4, 139-146, 1985
A;Title: Molecular cloning, sequencing, and expression in *Escherichia coli* of human prep
A;Reference number: I38102; MUID:85203359; PMID:3888571
A;Accession: I38102
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>
A;Cross-references: EMBL:X02760; NID:G35297; PIDN:CAA26535.1; PID:G35298
R;Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996
A;Title: Characterization of single chain urokinase-type plasminogen activator with a no
A;Reference number: S65783; MUID:96186279; PMID:8652631
A;Accession: S65783
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>
A;Cross-references: EMBL:D11143; NID:G1311467; PIDN:BAA01919.1; PID:G1199928
R;Gunzler, W.A.; Steffens, G.J.; Otting, F.; Küm, S.M.A.; Frankus, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
A;Title: The primary structure of high molecular mass urokinase from human urine.
A;Reference number: A37562; MUID:83055084; PMID:6754569
A;Accession: A37562
A;Molecule type: protein
A;Residues: 21-177 <GUN>
R;Schaller, J.; Nick, H.; Rickli, E.E.; Gillesse, D.; Lergier, W.; Studer, R.O. Eur. J. Biochem. 125, 251-257, 1982
A;Title: Human low-molecular-weight urinary urokinase. Partial characterization and prel
A;Reference number: A37563; MUID:83003608; PMID:6749491
A;Accession: A37563
A;Molecule type: protein
A;Residues: 156-176; 179-193, 'T', 195, 'T', 197-224 <SCH>
R;Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
A;Title: The complete amino acid sequence of low molecular mass urokinase from human uri
A;Reference number: A37564; MUID:83055099; PMID:6754572
A;Accession: A37564
A;Molecule type: protein
A;Residues: 158-410 <STE>
R;Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K. Biochem. Biophys. Res. Commun. 171, 401-406, 1990
A;Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinan
A;Reference number: A35689; MUID:90365737; PMID:23933398
A;Accession: A35689
A;Molecule type: protein
A;Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>
A;Note: Identification of a fucose and attempt to determine its attachment site
R;Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltzm Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
A;Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li
A;Reference number: A36697; MUID:91097529; PMID:2125213
A;Accession: A36697

A;Molecule type: protein
A;Residues: 21-34 <RAB>
R;Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M. submitted to the Brookhaven Protein Data Bank, July 1993
A;Cross-references: A51255; PDB:1KDU
A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
R;Li, X.; Smith, R.A.G.; Dobson, C.M. Biochemistry 31, 9562-9571, 1992
A;Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain f
A;Reference number: A44375; MUID:93003110; PMID:1327118
A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
R;Hansen, A.P.; Petros, A.M.; Meadows, R.F.; Netteheim, D.G.; Mazar, A.P.; Olejniczak, submitted to the Brookhaven Protein Data Bank, January 1994
A;Reference number: A66822; PDB:1URK
A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
R;Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.; submitted to the Brookhaven Protein Data Bank, July 1995
A;Reference number: A66058; PDB:1LMW
A;Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426
C;Comment: This enzyme is found in urine in a high molecular mass form, consisting of A
C;Genetics: Urokinase-type plasminogen activator proteolytically activates plasminogen, a
A;Gene: GDB:PLAU
A;Cross-references: GDB:119497; OMIM:191840
A;Map position: 10q24-10q24
A;Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
C;Function: proteolytically activates plasminogen
A;Pathway: fibrinolysis
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine proteina
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-431/Product: urokinase-type plasminogen activator, single chain form #status predic
F;21-177/Product: urokinase-type plasminogen activator chain A #status experimental <MPA
F;31-62/Domain: EGF homology <EGF>
F;70-151/Domain: kringle homology <KRG>
F;156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental <M
F;179-431/Product: urokinase-type plasminogen activator chain B #status experimental <MP
F;179-419/Domain: trypsin homology <TRY>
F;31-39 33-51 53-62 70-151 91-133 122-146 168-259 209-225 217-288 313-382 345-361 372-401
F;38/Binding site: carbohydrate (thr) (covalent) #status predicted
F;178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
F;224,275,376/Active site: His, Asp, Ser #status experimental
F;322/Binding site: carbohydrate (Asn) (covalent) #status experimental
Query Match 100.0%; Score 554; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.3e-51;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHN YCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHN YCRNPDN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 163
RESULT 2
UKBAY
u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
C;Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
C;Accession: S14687; S08651
R;Au, Y.P.T.; Wang, T.W.; Clowes, A.W. Nucleic Acids Res. 18, 3411, 1990
A;Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminoge
A;Reference number: S14687; MUID:90287734; PMID:2113276
A;Accession: S14687
A;Molecule type: mRNA
A;Residues: 1-433 <AU>
A;Cross-references: EMBL:X51935; NID:G38130; PIDN:CAA36200.1; PID:G38131
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-176/Product: plasminogen activator chain A #status predicted <ACH>
F;30-61/Domain: EGF homology <EGF>
F;59-150/Domain: kringle homology <KRG>
F;178-433/Product: plasminogen activator chain B #status predicted <BCH>
F;178-421/Domain: trypsin homology <TRY>
F;167-298,208-224,216-287,315-384,347-363,374-402/Disulfide bonds: #status predicted
F;223,274,378/Active site: His, Asp, Ser #status predicted
F;324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.7%; Score 519; DB 1; Length 433;
Best Local Similarity 94.8%; Pred. No. 7,4e-48; Indels 0; Gaps 0;
Matches 91; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLQOQTYHAHRSDALQLGLGKHYCRNPDN 60
DB 67 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLQOQTYHAHRSDALQLGLGKHYCRNPDN 126

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 96
DB 127 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 162

RESULT 3
UKPG
u-plasminogen activator (EC 3.4.21.73) precursor - pig
N;Alternate names: uPA
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
C;Accession: A00932
R;Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
Nucleic Acids Res. 12, 9526-9541, 1984
A;Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A;Reference number: A00932; MUID:85087954; PMID:6096832
A;Accession: A00932
A;Molecule type: DNA
A;Residues: 1-240, 'H', 242-442 <NAGI>
A;Experimental source: kidney cell line LLC-PK1
R;Nagamine, Y.
Submitted to the Protein Sequence Database, December 1986
A;Reference number: A37566
A;Contents: annotation; correction to residue 241
C;Genetics:
A;Introns: 19/3; 31/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;33-64/Domain: EGF homology <EGF>
F;72-153/Domain: kringle homology <KRG>
F;190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F;190-430/Domain: trypsin homology <TRY>
F;153/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;179-310,220-236,228-299,324-393,356-372,383-411/Disulfide bonds: #status predicted
F;235,286,387/Active site: His, Asp, Ser #status predicted

Query Match 79.0%; Score 437.5; DB 1; Length 442;
Best Local Similarity 74.3%; Pred. No. 4e-39;
Matches 78; Conservative 8; Mismatches 10; Indels 9; Gaps 1;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLQOQTYHAHRSDALQLGLGKHYCRNPDN 60
DB 70 QTCYEGNGHFYRGKASTDTMGRCPLPWNATVLQOQTYHAHRSDALQLGLGKHYCRNPDN 129

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 96
DB 130 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 174

RESULT 4
S18932
u-plasminogen activator (EC 3.4.21.73) precursor - rat

N;Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator
C;Species: Rattus norvegicus (Norway rat)
C;Date: 18-Oct-1989 #sequence_revision 10-Feb-1995 #text_change 18-Jun-1999
C;Accession: S24604; E160186; I53472; S18932
R;Rabbani, S.A.
Submitted to the EMBL Data Library, April 1992
A;Reference number: S24604
A;Accession: S24604
A;Molecule type: mRNA
A;Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>
A;Cross-references: EMBL:X65651; NID:G57456; PIDN:CAA46601.1; PID:G57457
A;Experimental source: tissue kidney
R;Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
Cancer Res. 52, 2489-2496, 1992
A;Title: Transcriptional and posttranscriptional activation of urokinase plasminogen ac
A;Reference number: E160186; MUID:92233409; PMID:1568219
A;Accession: I60186
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-432 <RES>
A;Cross-references: EMBL:X63434; NID:G57465; PIDN:CAA45028.1; PID:G57466
A;Experimental source: strain Fischer 344; tissue mammary
R;Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.
FEBS Lett. 306, 193-198, 1992
A;Title: The receptor for the plasminogen activator of urokinase type is up-regulated i
A;Reference number: I53472; MUID:92339549; PMID:1321734
A;Accession: I53472
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 31-62 <RE2>
A;Cross-references: EMBL:X66907; NID:G396200; PIDN:CAA47356.1; PID:G938279
C;Genetics:
A;Gene: uPA
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; tr
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;31-62/Domain: EGF homology <EGF>
F;70-151/Domain: kringle homology <KRG>
F;179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F;179-420/Domain: trypsin homology <TRY>
F;168-300,210-226,218-289,314-383,346-362,373-401/Disulfide bonds: #status predicted
F;225,276,377/Active site: His, Asp, Ser #status predicted

Query Match 77.1%; Score 427; DB 1; Length 432;
Best Local Similarity 77.1%; Pred. No. 5.1e-38;
Matches 74; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLQOQTYHAHRSDALQLGLGKHYCRNPDN 60
DB 68 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLQOQTYHAHRSDALQLGLGKHYCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 96
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 163

RESULT 5
UN0560
u-plasminogen activator (EC 3.4.21.73) precursor - bovine
N;Alternate names: uPA
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C;Accession: JN0560
R;Kraatzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A;Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and indu
A;Reference number: JN0560; MUID:93216119; PMID:8385052
A;Accession: JN0560
A;Molecule type: mRNA
A;Residues: 1-433 <KRA>
A;Cross-references: GB:I03546; NID:G163800; PIDN:AAA51419.1; PID:G163801
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; tr

C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-179/Product: plasminogen activator chain A #status predicted <MA1>
F;21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;33-64/Domain: EGF homology <EGF>
F;72-153/Domain: kringle homology <KRG>
F;181-433/Product: plasminogen activator chain B #status predicted <MA2>
F;181-421/Domain: trypsin homology <TRY>
F;170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F;226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 76.2%; Score 422; DB 1; Length 433;
Best Local Similarity 75.0%; Pred. No. 1.8e-37;
Matches 72; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

Oy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLOQTYHAHRSDALQLGLGKHNCRPN 60
Db 70 KTCYGNHGSYRGKANDLSGRCLANDSPVLLKWKYHAHRSDALQLGLGKHNCRPN 129

Oy 61 RRRPWCYVQGLKPLVQECWHDCAQDGKPSPPPEE 96
Db 130 QRRPWCYVQGLKPLVQECWHDCAQDGKPSPPPEE 165

RESULT 6

u-plasminogen activator (EC 3.4.21.73) precursor - mouse
C;Species: Mus.musculus (house mouse)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C;Accession: A29420; A24615
R;Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
Biochemistry 26, 8270-8279, 1987
A;Title: The murine urokinase-type plasminogen activator gene.
A;Reference number: A29420; MUID:85163489; PMID:281940
A;Accession: A29420
A;Molecule type: DNA
A;Residues: 1-433 <DEG>
A;Cross-references: GB:M17922; NID:G202296; PIDN:AAA40539.1; PID:G202297
R;Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher, Eur. J. Biochem. 148, 225-232, 1985
A;Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase
A;Reference number: A24615; MUID:85179474; PMID:2985383
A;Accession: A24615
A;Molecule type: mRNA
A;Residues: 1-433 <BEL>
A;Cross-references: GB:X02389; NID:G55127; PIDN:CAA26231.1; PID:G55128
C;Genetics:
A;Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;32-63/Domain: EGF homology <EGF>
F;71-152/Domain: kringle homology <KRG>
F;180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F;180-421/Domain: trypsin homology <TRY>
F;169-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F;226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 73.6%; Score 408; DB 1; Length 433;
Best Local Similarity 70.8%; Pred. No. 5.6e-36;
Matches 68; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

Oy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLOQTYHAHRSDALQLGLGKHNCRPN 60
Db 69 KTCYGNHGSYRGKANTDKGRCLANWAPVLOKPYNAHRSDALQLGLGKHNCRPN 128

Oy 61 RRRPWCYVQGLKPLVQECWHDCAQDGKPSPPPEE 96
Db 129 QRRPWCYVQGLKPLVQECWHDCAQDGKPSPPPEE 164

RESULT 7

J50599
t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: J50599
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boisdol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A;Reference number: J50597; MUID:92039036; PMID:1937019
A;Accession: J50599
A;Molecule type: mRNA
A;Residues: 1-431 <KRA>
A;Cross-references: GB:M3989; NID:G166076; PIDN:AAA31594.1; PID:G166077
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-431/Product: plasminogen activator beta #status predicted <PLA>
F;41-74/Domain: EGF homology <EGF>
F;82-163/Domain: kringle homology <KRG>
F;180-425/Domain: trypsin homology <TRY>
F;41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide bon
F;139,352/Binding site: carboxylate (Asn) (covalent) #status predicted
F;179-180/Cleavage site: His-Ser (plasmin) #status predicted
F;226,275,382/Active site: His, Asp, Ser #status predicted
F;345-361,378-406/Disulfide bonds: #status predicted

Query Match 43.5%; Score 241; DB 2; Length 431;
Best Local Similarity 50.0%; Pred. No. 4.1e-18;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

Oy 2 TCYEGNGHFYRGKASTDTMGRCPLPWNATVLOQTYHAHRSDALQLGLGKHNCRPN 61
Db 81 TCYKQGVYTRGTWSTSSGACINWNSNLLTRTYNGRSDATLGLGKHNCRPN 140

Oy 62 RRPWCYVQGLKPLVQECWHDCA 85
Db 141 SKPWCYVQGLKPLVQECWHDCA 164

RESULT 8

A34369
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
C;Species: Megaderma lyra
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A34369
R;Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob
J. Biol. Chem. 264, 17947-17952, 1989
A;Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin
A;Reference number: A34369; MUID:90036867; PMID:2509450
A;Accession: A34369
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-477 <GAR>
A;Cross-references: GB:J05082; NID:G166080; PIDN:AAA31596.1; PID:G166081
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-477/Product: plasminogen activator #status predicted <PLA>
F;42-79/Domain: fibronectin type I repeat homology <IFA>
F;87-120/Domain: EGF homology <EGF>
F;128-209/Domain: kringle homology <KRG>
F;226-471/Domain: trypsin homology <TRY>
F;42-72,70-79,87-98,92-105,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 43.5%; Score 241; DB 1; Length 477;
Best Local Similarity 50.0%; Pred. No. 4.5e-18;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

Oy 2 TCYEGNGHFYRGKASTDTMGRCPLPWNATVLOQTYHAHRSDALQLGLGKHNCRPN 61

Db 186 SKPCYYFKAGKYSSEFCSTPACSEG 211

RESULT 12

UKHUT

t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
 N:Alternate names: t-PA; tissue plasminogen activator
 C:Species: Homo sapiens (man)
 C>Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 08-Dec-2000
 C:Accession: A94004; A23529; J0562; A93293; S02125; A91343; A93951; A91322; A54645; 160
 R:NY, T.; Bligh, P.; Lund, B.
 PROC. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
 A:Title: The structure of the human tissue-type plasminogen activator gene: correlation
 A:Reference number: A94004; MUID:84298137; PMID:6089199
 A:Accession: A94004
 A:Molecule type: DNA
 A:Residues: 1-562 <NYT>
 A:Cross-references: GB:L00141
 A:Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translation
 R:Friezner Degen, S.J.; Rajput, B.; Reich, E.
 J. Biol. Chem. 261, 6972-6985, 1986
 A:Title: The human tissue plasminogen activator gene.
 A:Reference number: A23529; MUID:86196143; PMID:3009482
 A:Accession: A23529
 A:Molecule type: DNA
 A:Residues: 1-562 <DEG>
 A:Cross-references: GB:K03021; NID:G339817; PIDN:AA98809.1; PID:G339818
 R:Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuoka, S.; Higashio, K.
 Agric. Biol. Chem. 55, 1225-1232, 1991
 A:Title: Purification and characterization of tissue plasminogen activator secreted by h
 A:Reference number: J0562; MUID:91291340; PMID:1368681
 A:Accession: J0562
 A:Molecule type: mRNA
 A:Residues: 31-562 <ITA>
 A:Cross-references: DBJL:D01096; NID:9220128; PIDN:BA000881.1; PID:G441174
 A:Experimental source: embryonic lung fibroblast IMR-90 cells
 R:Penica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vehar, C.A.; Ward, C.A.; Bennett
 Nature 301, 214-221, 1983
 A:Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
 A:Reference number: A93293; MUID:83115262; PMID:6337343
 A:Accession: A93293
 A:Molecule type: mRNA
 A:Residues: 1-562 <PEN>
 A:Cross-references: GB:L00141
 A:Experimental source: melanoma cells
 R:Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
 Nucleic Acids Res. 15, 5695, 1988
 A:Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
 A:Reference number: S02125; MUID:85285279; PMID:3133640
 A:Accession: S02125
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-562 <SAS>
 A:Cross-references: EMBL:X07393; NID:G37243; PIDN:CAA30302.1; PID:G37244
 A:Experimental source: fetal lung cells
 R:Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma
 FEBS Lett. 189, 145-149, 1985
 A:Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
 A:Reference number: A91343; MUID:85285620; PMID:3896853
 A:Accession: A91343
 A:Molecule type: mRNA
 A:Residues: 1-38, 'G', 86-433, 'E', 435-562 <KAG>
 A:Experimental source: Detroit 562 cells; ATCC 138
 R:Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
 Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
 A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activ
 A:Reference number: A93951; MUID:83169656; PMID:6572897
 A:Accession: A93951
 A:Molecule type: mRNA
 A:Residues: 251-358 <BDL>
 A:Experimental source: melanoma cells
 R:Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.

Biochemistry 23, 3701-3707, 1984
 A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
 differences.
 A:Reference number: A90488; MUID:85000468; PMID:6433976
 A:Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and I
 R:Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.
 FEBS Lett. 168, 29-32, 1984
 A:Title: Differences between uterine and melanoma forms of tissue plasminogen activator.
 A:Reference number: A91322; MUID:84158956; PMID:6538514
 A:Accession: A91322
 A:Molecule type: protein
 A:Residues: 33-45; 311-320 <POH>
 A:Experimental source: uterus
 A:Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
 R:Van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
 J. Biol. Chem. 261, 14214-14218, 1986
 A:Reference number: A37567; MUID:87033611; PMID:3021732
 R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Enger
 EXBO J. 5, 3525-3530, 1986
 A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen ac
 A:Reference number: A37568; MUID:87161761; PMID:3030730
 A:Contents: annotation; fibrin binding site
 R:Dodd, I.; Nunn, B.; Robinson, J.H.
 Thromb. Haemost. 59, 523-528, 1988
 A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type p
 A:Reference number: A60902; MUID:89044681; PMID:3142086
 A:Contents: annotation; novel forms of expressed recombinant t-PA
 R:Haris, T.J.R.; Patel, T.; Maston, F.A.O.; Lattie, S.; Emtage, J.S.; Opdenakker, G.;
 Mol. Biol. Med. 3, 279-292, 1986
 A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expr
 A:Reference number: A54645; MUID:86284200; PMID:3090401
 A:Accession: A54645
 A:Molecule type: mRNA
 A:Residues: 1-562 <HAR>
 A:Cross-references: GB:M5518; NID:G190031; PIDN:AAA60111.1; PID:G190032
 R:Reddy, V.B.; Garrazone, A.J.; Sasaki, H.; Wei, C.
 DNA 6, 461-472, 1987
 A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells us
 A:Reference number: I60110; MUID:88054470; PMID:2824147
 A:Accession: I60110
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-562 <RES>
 A:Cross-references: GB:M18182; NID:G340176; PIDN:AAA36800.1; PID:G340177
 R:Fishel, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
 J. Biol. Chem. 260, 11223-11230, 1985
 A:Title: Isolation and characterization of the human tissue-type plasminogen activator s
 A:Reference number: I55232; MUID:85289338; PMID:3161893
 A:Accession: I55232
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-36 <RE2>
 A:Cross-references: GB:M1890; NID:G339837; PIDN:AAA61213.1; PID:G339839
 C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single
 C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond. It
 C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat.
 C:Genetics:
 A:Gene: GDB:PLAT
 A:Cross-references: GDB:119496; OMIM:173370
 A:Map position: 8p12-8p12
 A:Introns: 24/3, 39/1, 85/1, 122/1, 180/2, 211/1, 268/2, 297/1, 362/2, 408/1, 455/3, 510
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-32/Domain: propeptide #status predicted <PRO>
 F:33-562/Product: t-plasminogen activator #status experimental <MAT>
 F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
 F:41-78/Domain: fibronectin type I repeat homology <IFI>
 F:86-119/Domain: EGF homology <EGF>
 F:127-208/Domain: kringle homology <KRI>
 F:215-296/Domain: kringle homology <KR2>

F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
 F:311-556/Domains: trypsin homology <TRY>
 F:41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203,215-286,236-278,267-291,299-4
 F:152,483/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:219/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
 F:310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental
 F:357,405/Active site: His, Asp #status predicted
 F:513/Active site: Ser #status experimental

Query Match 40.8%; Score 226; DB 1; Length 562;
 Best Local Similarity 47.7%; Pred. No. 2.1e-16;
 Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCVEGNHFRYKASTDTMGRLCPWNSATVLQOYTHAHRSDALQGLGKHNCRPNDR 61
 126 TCYEDQGISYRTWTSTASGAECTWNSSALAAQKPYSGRRDARLGLGNHNYCENPDRD 185

Db

QY 62 RRPWCYVQGLKPLVQECMVHDCADG 87
 186 SRPWCYVFKAGKYSFECSTPACSEG 211

Db

RESULT 13

t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
 N:Alternate names: tissue plasminogen activator
 C:Species: Desmodus rotundus (common vampire bat)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C:Accession: J50600
 R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
 Gene 105, 229-237, 1991
 A:Title: The plasminogen activator family from the salivary gland of the vampire bat Des
 A:Reference number: JS0597; MUID:92039036; PMID:1937019
 A:Accession: J50600

A:Molecule type: mRNA
 A:Residues: 1-394 <KRA>
 A:Cross-references: GB:M63990; NID:q166078; PIDN:AAA31595.1; PID:G166079
 A:Note: the authors translated the codon ATC for residue 75 as Thr
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F:1-21/Domains: signal sequence #status predicted <SIG>
 F:22-36/Domains: propeptide #status predicted <PRO>
 F:37-394/Product: plasminogen activator gamma #status predicted <PLA>
 F:45-126/Domains: kringle homology <KEG>
 F:143-388/Domains: trypsin homology <TRY>
 F:45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:
 F:142-143/Cleavage site: His-Ser (plasmin) #status predicted
 F:189,238,345/Active site: His, Asp, Ser #status predicted
 F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 39.7%; Score 220; DB 2; Length 394;
 Best Local Similarity 45.2%; Pred. No. 6.6e-16;
 Matches 38; Conservative 13; Mismatches 33; Indels 0; Gaps 0;

QY 2 TCVEGNHFRYKASTDTMGRLCPWNSATVLQOYTHAHRSDALQGLGKHNCRPNDR 61
 44 TCYKQGVYRTWTSTBSGAQCNWNSLLIRRTYNGRMPVAVKLGNGHNYCENPDGA 103

Db

QY 62 RRPWCYVQGLKPLVQECMVHDCA 85
 104 SRPWCYVFKARKFTSESCVPVCS 127

Db

RESULT 14

t-plasminogen activator (EC 3.4.21.68) precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A29941; S48205; S48207; S48206
 R:Rickles, R.J.; Darrow, A.L.; Strickland, S.
 J. Biol. Chem. 263, 1563-1569, 1988
 A:Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator m
 A:Reference number: A29941; MUID:88087303; PMID:2826484

A:Accession: A29941
 A:Molecule type: mRNA
 A:Residues: 1-559 <RIC>
 A:Cross-references: GB:J09520; NID:q202109; PIDN:AAA04070.1; PID:q202110
 R:Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
 Eur. J. Biochem. 224, 863-871, 1994
 A:Title: Characterization of the murine plasma fibrinolytic system.
 A:Reference number: S48202; MUID:95010076; PMID:7523120
 A:Accession: S48205

A:Molecule type: protein
 A:Residues: 33-37, 'X', 39-40 <LIJ>
 A:Accession: S48207
 A:Molecule type: protein
 A:Residues: 309-316 <L12>
 A:Accession: S48206
 A:Molecule type: protein
 A:Residues: 33-37, 'X', 39-40 <LIW>
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F:1-17/Domains: signal sequence #status predicted <SIG>
 F:18-29/Domains: propeptide #status predicted <PRO>
 F:30-559/Product: t-plasminogen activator #status predicted <MAT>
 F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
 F:38-75/Domains: fibronectin type I repeat homology <IF1>
 F:83-116/Domains: EGF homology <EGF>
 F:124-205/Domains: kringle homology <KR1>
 F:213-294/Domains: kringle homology <KR2>
 F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
 F:309-553/Domains: trypsin homology <TRY>
 F:38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-
 F:149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
 F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 39.5%; Score 219; DB 1; Length 559;
 Best Local Similarity 46.0%; Pred. No. 1.2e-15;
 Matches 40; Conservative 11; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCVEGNHFRYKASTDTMGRLCPWNSATVLQOYTHAHRSDALQGLGKHNCRPNDR 61
 123 TCVEQGITVYRTWTSTASGAECTWNSSVLKPYNARRPNAIKLGLGNHNYCENPDRD 182

Db

QY 62 RRPWCYVQGLKPLVQECMVHDCADGK 88
 183 LKWCYVFKAGKVTTFECSTPACPKK 209

Db

RESULT 15

JS0597

t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
 N:Alternate names: tissue plasminogen activator
 C:Species: Desmodus rotundus (common vampire bat)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C:Accession: JS0597
 R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Do
 Gene 105, 229-237, 1991
 A:Title: The plasminogen activator family from the salivary gland of the vampire bat De
 A:Reference number: JS0597; MUID:92039036; PMID:1937019
 A:Accession: JS0597

A:Molecule type: mRNA
 A:Residues: 1-477 <KRA>
 A:Cross-references: GB:M63987; NID:G166070; PIDN:AAA31591.1; PID:G166071
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F:1-21/Domains: signal sequence #status predicted <SIG>
 F:22-36/Domains: propeptide #status predicted <PRO>
 F:37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>
 F:42-79/Domains: fibronectin type I repeat homology <IF1>
 F:87-120/Domains: EGF homology <EGF>
 F:128-209/Domains: kringle homology <KRG>
 F:226-471/Domains: trypsin homology <TRY>
 F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-
 F:153,358/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:225-226/Cleavage site: His, Ser (plasmin) #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 38.4%; Score 213; DB 2; Length 477;
Best Local Similarity 46.4%; Pred. No. 4.4e-15;
Matches 39; Conservative 10; Mismatches 35; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNCRPNRR 61
DB 127 TCYEGGQVTVRGTWSTAESRVEICINWNSLLTRTYNGRMPDAFNLGLGNHNYCRPNRGA 186

QY 62 RRPWCYVQGLKPLVQECMVHDC 85

DB 187 RRPWCYVQGLKPLVQECMVHDC 210

RESULT 16

JC5878

Plasma hyaluronan-binding protein precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999

C:Accession: JC5878

R:Hashimoto, K.; Tobe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, N.

Biol. Pharm. Bull. 20, 1127-1130, 1997

A:Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-binding protein

A:Reference number: JC5878; MUID:98045239; PMID:9401717

A:Accession: JC5878

A:Molecule type: mRNA

A:Residues: 1-558 <HAS>

C:Comment: This protein acts as serine protease.

C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; tryptophan

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <MATL>

F:75-106/Domain: EGF homology <EG1>

F:113-145/Domain: EGF homology <EG2>

F:152-185/Domain: EGF homology <EG3>

F:192-274/Domain: kringle homology <KRI>

F:312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MATL>

F:312-548/Domain: trypsin homology <TRY>

Query Match 37.9%; Score 210; DB 2; Length 558;

Best Local Similarity 43.5%; Pred. No. 1.1e-14;

Matches 40; Conservative 14; Mismatches 34; Indels 4; Gaps 2;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNCRPNRR 62

DB 192 CYVGDGYSYRGKSKVTNQNPCLYWNHLLQETYNMFEDAEHTGIAEHFPCRNPDGDH 251

QY 63 RRPWCYVQGLKPLVQECMVHDCADGKPS 93

DB 252 RRPWCYVQGLKPLVQECMVHDCADGKPS 280

RESULT 17

JC4795

Plasma hyaluronan-binding protein precursor - human

N:Alternate names: hepatocyte growth factor activator-like protein; PHBP

N:Contains: serine proteinase (EC 3.4.21.-)

C:Species: Homo sapiens (man)

C:Date: 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 19-Jul-2002

C:Accession: JC4795

R:Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, T.; Tomita, M.

J. Biochem. 119, 1157-1165, 1995

A:Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP)

x activator.

A:Reference number: JC4795; MUID:96425001; PMID:8827452

A:Accession: JC4795

A:Molecule type: mRNA

A:Residues: 1-560 <CHO>

A:Cross-references: GB:S83182; NID:g1836158; PIDN:RAB46909.1; PID:g1836159

A:Experimental source: plasma

A>Note: parts of this sequence, including the amino ends of the mature chains, were detected

C:Genetics:

A:Gene: GDB:HABP2; HABP; PHBP; HGFAL

A:Cross-references: GDB:4573962

C:Complex: a disulfide-bonded heterodimer of chains produced from the same precursor; the

C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; tryptophan

C:Keywords: Chondroitin sulfate proteoglycan; glycoprotein; hyaluronic acid; hydrolase; 1

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-311/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>

F:77-108/Domain: EGF homology <EG1>

F:115-147/Domain: EGF homology <EG2>

F:154-187/Domain: EGF homology <EG3>

F:194-276/Domain: kringle homology <KRI>

F:314-550/Domain: trypsin homology <TRY>

F:314-516/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted

F:54,207/Binding site: carboxylate (Asn) (covalent) #status predicted

F:77-88,82-97,99-108,115-125,120-136,138-147,154-165,159-176,178-187,194-276,215-257,246-

F:362,405,509/Active site: His, Asp, Ser #status predicted

Query Match 36.0%; Score 199.5; DB 1; Length 560;

Best Local Similarity 40.2%; Pred. No. 1.4e-13;

Matches 39; Conservative 15; Mismatches 40; Indels 3; Gaps 2;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNCRPNRR 62

DB 194 CYVGDGYSYRGKSKVTNQNPCLYWNHLLQETYNMFEDAEHTGIAEHFPCRNPD 253

QY 63 RRPWCYVQGLKPLVQECMVHDC--DGKKPSSPPEE 96

DB 254 RRPWCYVQGLKPLVQECMVHDC--DGKKPSSPPEE 290

RESULT 18

A35005

U-plasminogen activator (EC 3.4.21.73) precursor - chicken

N:Alternate names: uPA

C:Species: Gallus gallus (chicken)

C:Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999

C:Accession: A35005

J. Biol. Chem. 265, 1339-1344, 1990

A:Title: The chicken urokinase-type plasminogen activator gene.

A:Reference number: A35005; MUID:90110185; PMID:2295632

A:Accession: A35005

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-434 <UES>

A:Cross-references: GB:J05187; NID:g212858; PIDN:AAA49131.1; PID:g212859

C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; tryptophan

C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>

F:40-71/Domain: EGF homology <EGF>

F:79-158/Domain: kringle homology <KRG>

F:173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>

F:173-416/Domain: trypsin homology <TRY>

F:162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted

F:217,272,373/Active site: His, Asp, Ser #status predicted

Query Match 35.8%; Score 198.5; DB 1; Length 434;

Best Local Similarity 54.4%; Pred. No. 1.4e-13;

Matches 37; Conservative 7; Mismatches 19; Indels 5; Gaps 2;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQ-QTYHAHRSDALQGLGKHNCRPNRR 61

DB 79 CYSGNGDYRGNAEDP---GCLYWDHPSVIRMGDYHADLNKALQGLGKHNCRPNRR 134

QY 62 RRPWCYVQ 69

DB 135 RRPWCYVQ 142

RESULT 19

A4688

hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human

C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
C;Accession: A46688
R;Wiyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
J. Biol. Chem. 268, 10024-10028, 1993
A;Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease d coagulation factor XII.
A;Reference number: A46688; MUID:93252878; PMID:7683665
A;Accession: A46688
A;Molecule type: mRNA
A;Residues: 1-655 <NIV>
A;Cross-references: DDBJ:D14012; NID:g219680; PID:BAA03113.1; PID:g219681
A;Experimental source: Liver (mRNA); serum (protein)
A;Note: sequence extracted from NCBI backbone (NCBIN:131227, NCBI:P:131228)
A;Note: parts of the sequence, including the amino ends of the heavy and light chains, c C;Genetics:
A;Gene: GDB:HGFA; HGFA; HGFA
A;Cross-references: GDB:9954514
A;Map position: 4p16-4p16
C;Function:
A;Description: activates hepatocyte growth factor by specific proteolytic cleavage
A;Pathway: tissue repair and regeneration
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C;Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
F;1-34/Domain: signal sequence #status predicted <SIG>
F;108-148/Domain: fibronectin type II repeat homology <IF2>
F;164-197/Domain: EGF homology <EG1>
F;202-237/Domain: fibronectin type I repeat homology <IF1>
F;245-278/Domain: EGF homology <EG2>
F;286-367/Domain: kringle homology <KRG>
F;373-407/Product: hepatocyte growth factor activator light chain #status experimental
F;408-555/Product: hepatocyte growth factor activator heavy chain #status experimental
F;408-541/Domain: trypsin homology <TRY>
F;40,48,290,468,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;164-175,169-186,188-197,202-237,245-256,250-267,269-278,286-367,307-349,338-36 F;447,497,598/Active site: His, Asp, Ser #status predicted
Query Match 35.1%; Score 194.5; DB 1; Length 655;
Best Local Similarity 40.4%; Pred. No. 5.7e-13;
Matches 42; Conservative 8; Mismatches 41; Indels 13; Gaps 1;
QY 3 CYEGNGHFYRGKASTDGMORCLPWN SATVLOQTYHAHRS DALQGLGKHNYCRNPDRR 62
DB 286 CFLGNGTGYRGVASTASGLSCLAWNSDLLYQELHVDVSGAALLGLGPHAYCRNPDRNDE 345
QY 63 RPKCYVQVGLKPLVQSCVHDC-----ADGKKPSP 93
DB 346 RPKCYVQVGLKPLVQSCVHDC-----ADGKKPSP 389
RESULT 20
coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)
N;Alternate names: Hageman factor
C;Species: Cavia porcellus (guinea pig)
C;Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 21-Jan-2000
C;Accession: S28941
R;Samba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambata, T.; Okabe, H. Biochim. Biophys. Acta 1159, 113-121, 1992
A;Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage si A;Reference number: S28941; MUID:93003367; PMID:1390917
A;Accession: S28941
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-603 <SEM>
A;Cross-references: EMBL:X68615; NID:g49578; PIDN:CAA48600.1; PID:g49579
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C;Keywords: hydrolase; serine proteinase
F;46-87/Domain: fibronectin type II repeat homology <IF2>
F;134-169/Domain: fibronectin type I repeat homology <FBI>
F;177-208/Domain: EGF homology <EGF>
F;216-294/Domain: kringle homology <KRG>
F;359-597/Domain: trypsin homology <TRY>

Query Match 34.6%; Score 191.5; DB 2; Length 603;
Best Local Similarity 39.6%; Pred. No. 1.1e-12;
Matches 38; Conservative 15; Mismatches 38; Indels 5; Gaps 2;
QY 2 TCYEGNGHFYRGKASTDGMORCLPWN SATVLOQTYHAHRS DALQGLGKHNYCRNPDRN 60
DB 215 SCYEGRGVSRYGMARTVTSQAKQWRAS----BATYRNMTAEQALRRGLGHTFCRNPDRN 270
QY 61 RRRPCCYVQVGLKPLVQSCVHDCADGKKPSPSP 96
DB 271 DTRPCCVFWNGNRLSNWEYCDLAQCQYPPQPTATPHD 306
RESULT 21
coagulation factor XIIa (EC 3.4.21.38) precursor [validated] - human
N;Alternate names: Hageman factor (activated)
C;Species: Homo sapiens (man)
C;Date: 27-Nov-1985 #sequence_revision 30-Jun-1991 #text_change 08-Dec-2000
C;Accession: A29411; A26814; A00930; A25191; A22248; A21037
R;Cool, D.E.; MacGillivray, R.T.A.
J. Biol. Chem. 262, 13662-13673, 1987
A;Title: Characterization of the human blood coagulation factor XII gene. Intron/exon 9 A;Reference number: A29411; MUID:88007593; PMID:2888762
A;Accession: A29411
A;Molecule type: DNA
A;Residues: 1-615 <COO>
A;Cross-references: GB:M17466; GB:J02807; NID:g180355; PIDN:AAB59490.1; PID:g180357
R;Tripiodi, M.; Citarella, F.; Guida, S.; Galeffi, P.; Fantoni, A.; Cortese, R. Nucleic Acids Res. 14, 3146, 1986
A;Title: cDNA sequence coding for human coagulation factor XII (Hageman).
A;Reference number: A26814; MUID:86176794; PMID:3754331
A;Accession: A26814
A;Molecule type: mRNA
A;Residues: 4-615 <TRI>
A;Cross-references: GB:M31315; NID:g182291; PIDN:AAA70225.1; PID:g182292
R;Cool, D.E.; Edgell, C.J.S.; Louie, G.V.; Zoller, M.J.; Brayer, G.D.; MacGillivray, R. J. Biol. Chem. 260, 13666-13675, 1985
A;Title: Characterization of human blood coagulation factor XII cDNA. Prediction of the A;Reference number: A00930; MUID:86033830; PMID:3877053
A;Accession: A00930
A;Molecule type: mRNA
A;Residues: 14-332, 'S', 334-615 <CO2>
R;Que, B.G.; Davis, E.W. Biochemistry 25, 1523-1528, 1986
A;Title: Characterization of a cDNA coding for human factor XII (Hageman factor).
A;Reference number: A25191; MUID:86216049; PMID:3011063
A;Accession: A25191
A;Molecule type: mRNA
A;Residues: 146-378, 'G', 380-615 <QUE>
A;Cross-references: GB:M3147; NID:g180360; PIDN:AAA70224.1; PID:g180361
R;McMullen, B.A.; Fujikawa, K. J. Biol. Chem. 260, 5328-5341, 1985
A;Title: Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated H A;Reference number: A22248; MUID:85182674; PMID:3886654
A;Accession: A22248
A;Molecule type: protein
A;Residues: 20-379 <WCM>
R;Fujikawa, K.; McMullen, B.A. J. Biol. Chem. 258, 10924-10933, 1983
A;Title: Amino acid sequence of human beta-factor XIIa.
A;Reference number: A21037; MUID:83291041; PMID:6604055
A;Accession: A21037
A;Molecule type: protein
A;Residues: 354-362; 373-615 <FUI>
R;Harris, R.J.; Lang, V.T.; Spellman, M.W. J. Biol. Chem. 267, 5102-5107, 1992
A;Title: O-linked fucose is present in the first epidermal growth factor domain of fact A;Reference number: A44606; MUID:92184750; PMID:1544894
A;Contents: annotation; carbohydrate binding site
C;Genetics:

A;Gene: GDB:F12
A;Cross-references: GDB:119892; OMIM:234000
A;Map position: Sq34-5qter
A;Introns: 19/3; 72/2; 96/1; 133/1; 177/1; 212/1; 267/2; 340/1; 417/2; 463/1; 511/1
A;Complex: factor XII, prekallikrein, and HNW kininogen form a complex bound to anionic C;Function:
A;Description: factor XIIa catalyzes the proteolytic activation of plasminogen, plasma prekallikrein
A;Pathway: blood coagulation; fibrinolysis
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology; kringle
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; kringle; plasma; protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-372,373-615/Product: coagulation factor XIIa, alpha form #status experimental <A12>
F;47-88/Domain: fibronectin type II repeat homology <FB2>
F;98-130/Domain: EGF homology <EG1>
F;135-170/Domain: fibronectin type I repeat homology <1F1>
F;178-209/Domain: EGF homology <EG2>
F;217-295/Domain: kringle homology <KEG>
F;298-356/Region: proline-rich
F;334-362,373-615/Product: coagulation factor XIIa, beta form #status experimental <B12>
F;373-609/Domain: trypsin homology <TRY>
F;98-110,104-119,121-130,135-163,161-170,178-189,193-198,200-209,217-295,238-277,266-290
F;109/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;249,433/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;299,305,328,329,337/Binding site: carbohydrate (Thr) (covalent) #status predicted
F;308/Binding site: carbohydrate (Ser) (covalent) #status predicted
F;412,461,563/Active site: His, Asp, Ser #status predicted
Query Match 30.8%; Score 170.5; DB 1; Length 615;
Best Local Similarity 48.5%; Pred. No. 2e-10;
Matches 33; Conservative 6; Mismatches 24; Indels 5; Gaps 2;
QY 2 TCYEGNGHYRGKASTDTMGPRCLPWN SATVLQTY-HAHRSDALQLGLGKHYCRNP DN 60
Db 216 SCYDGRGLSYRGARTLTLSGAPCPQWAS-----EATYRNTVAEQARNWGLGHPACFRNP DN 271
QY 61 RRRPWCYV 68
Db 272 DIRPWCYV 279
RESULT 22
A32869
Apolipoprotein(a) (EC 3.4.21.-) - rhesus macaque (fragment)
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 22-Jun-1999
C;Accession: A32869; A30848
R;Tomlinson, J.E.; McLean, J.W.; Lamn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A;Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A;Reference number: A32869; MUID:89174660; PMID:2925643
A;Accession: A32869
A;Molecule type: mRNA
A;Residues: 1-1420 <TOM>
A;Cross-references: GB:J04635; NID:G342072; PIDN:AAA36833.1; PID:G342073
C;Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C;Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F;50-127/Domain: kringle homology <KR1>
F;164-241/Domain: kringle homology <KR2>
F;278-355/Domain: kringle homology <KR3>
F;392-469/Domain: kringle homology <KR4>
F;506-583/Domain: kringle homology <KR5>
F;620-697/Domain: kringle homology <KR6>
F;726-803/Domain: kringle homology <KR7>
F;840-917/Domain: kringle homology <KR8>
F;954-1031/Domain: kringle homology <KR9>
F;1068-1145/Domain: kringle homology <KR10>
F;1191-1413/Domain: trypsin homology <TRY>
Query Match 29.1%; Score 161; DB 2; Length 1420;
Best Local Similarity 40.9%; Pred. No. 4.6e-09;
Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;

QY 3 CYEGNGHYRGKASTDTMGPRCLPWN SATVLQ--QTYHAHRSDALQLGLGKHYCRNP DN 60
Db 1068 CYHNGQSYRTFTTGTTCQSNWSTPHQKRTTENTPNDLTM-----NYCRNPDA 1122
QY 61 RRRPWCYVQVGLKPLVQE--CMWHDCA 86
Db 1123 DTGPWCFT---MDPSVRREYCNLTRCSD 1147
RESULT 23
S00657
Apolipoprotein(a) (EC 3.4.21.-) precursor [validated] - human
N;Alternate names: apolipoprotein(a); lipoprotein(a) chain apo(a)
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 08-Dec-2000
C;Accession: S00657; A28017; A47277; I50906; A47233; I52415; I65286
R;McLean, J.W.; Tomlinson, J.E.; Kuang, W.J.; Eaton, D.L.; Chen, E.Y.; Fleess, G.M.; Scan
Nature 330, 132-137, 1987
A;Title: cDNA sequence of human apolipoprotein(a) is homologous to plasminogen.
A;Reference number: S00657; MUID:88039109; PMID:3670400
A;Accession: S00657
A;Molecule type: mRNA
A;Residues: 1-4548 <MCL>
A;Cross-references: GB:X06290; EMBL:X06696; NID:G28619; PIDN:CAA29618.1; PID:G28620
R;Eaton, D.L.; Fleess, G.M.; Kohr, W.J.; McLean, J.W.; Xu, Q.T.; Miller, C.G.; Lamn, R.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 3224-3228, 1987
A;Title: Partial amino acid sequence of apolipoprotein(a) shows that it is homologous to
A;Reference number: A28017; MUID:87204109; PMID:3472206
A;Accession: A28017
A;Molecule type: protein
A;Residues: 20-21, P', 23-34, 177-179, N', 181-186, T', 188-196, DKG', 200; 292-314, W', 316-31
X', 4336-4401 <EAT>
R;Wade, D.P.; Clarke, J.G.; Lindahl, G.E.; Liu, A.C.; Zysow, B.R.; Meer, K.; Schwartz, K
Proc. Natl. Acad. Sci. U.S.A. 90, 1369-1373, 1993
A;Title: 5' control regions of the apolipoprotein(a) gene and members of the related pla
A;Reference number: A47277; MUID:93165698; PMID:7679504
A;Accession: A47277
A;Status: preliminary; translation not shown; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-16 <RES>
A;Cross-references: GB:L07899; NID:G967973; PID:G967974
R;Malgaretti, N.; Acquati, F.; Magnaghi, P.; Bruno, L.; Pontoglio, M.; Rocchi, M.; Sacco
Proc. Natl. Acad. Sci. U.S.A. 89, 11584-11588, 1992
A;Title: Characterization by yeast artificial chromosome cloning of the linked apolipoppr
A;Reference number: A47233; MUID:93087573; PMID:1454851
A;Accession: I60906
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-16 <RE2>
A;Cross-references: GB:M90078; NID:G178786; PIDN:AAA35547.1; PID:G553188
A;Note: apo(a) gene 1 (nomenclature of reference I52415)
A;Accession: A47233
A;Status: preliminary; translation not shown; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-16 <RE5>
A;Cross-references: GB:M90079; NID:G178784; PIDN:AAA35546.1; PID:G553187
R;Ichinose, A.
Biochemistry 31, 3113-3118, 1992
A;Title: Multiple members of the plasminogen-apolipoprotein(a) gene family associated wi
A;Reference number: I52415; MUID:92207924; PMID:1554698
A;Accession: I52415
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-16 <RE3>
A;Cross-references: GB:M86877; NID:G178780; PIDN:AAA49909.1; PID:G553185
A;Note: apo(a) gene 1 (nomenclature of reference I52415)
A;Accession: I65286
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-16 <RE4>
A;Cross-references: GB:M86878; NID:G178782; PIDN:AAA51749.1; PID:G553186
C;Genetics:
A;Gene: GDB:LPA

A;Cross-references: GDB:120699; OMIM:152200
A;Map position: 6q26-6q27
A;Note: several genes closely linked on chromosome 6 are identical in the first coding regions of kringle repeats
C;Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C;Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-4548/Product: apolipoprotein(a) #status experimental <MAT>
F;142-219/Domain: kringle homology <KR1>
F;256-333/Domain: kringle homology <KR2>
F;370-447/Domain: kringle homology <KR3>
F;484-561/Domain: kringle homology <KR4>
F;588-675/Domain: kringle homology <KR5>
F;712-789/Domain: kringle homology <KR6>
F;826-903/Domain: kringle homology <KR7>
F;940-1017/Domain: kringle homology <KR8>
F;1054-1131/Domain: kringle homology <KR9>
F;1168-1245/Domain: kringle homology <KR10>
F;1282-1359/Domain: kringle homology <KR11>
F;1396-1473/Domain: kringle homology <KR12>
F;1510-1587/Domain: kringle homology <KR13>
F;1624-1701/Domain: kringle homology <KR14>
F;1738-1815/Domain: kringle homology <KR15>
F;1852-1929/Domain: kringle homology <KR16>
F;1966-2043/Domain: kringle homology <KR17>
F;2080-2157/Domain: kringle homology <KR18>
F;2194-2271/Domain: kringle homology <KR19>
F;2308-2385/Domain: kringle homology <KR20>
F;2422-2499/Domain: kringle homology <KR21>
F;2536-2613/Domain: kringle homology <KR22>
F;2650-2727/Domain: kringle homology <KR23>
F;2764-2841/Domain: kringle homology <KR24>
F;2878-2955/Domain: kringle homology <KR25>
F;2992-3069/Domain: kringle homology <KR26>
F;3106-3183/Domain: kringle homology <KR27>
F;3202-3297/Domain: kringle homology <KR28>
F;3334-3411/Domain: kringle homology <KR29>
F;3448-3525/Domain: kringle homology <KR30>
F;3562-3639/Domain: kringle homology <KR31>
F;3676-3753/Domain: kringle homology <KR32>
F;3782-3859/Domain: kringle homology <KR33>
F;3896-3973/Domain: kringle homology <KR34>
F;4010-4087/Domain: kringle homology <KR35>
F;4124-4201/Domain: kringle homology <KR36>
F;4228-4307/Domain: kringle homology <KR37>
F;4328-4541/Domain: trypsin homology <TRY>

Query Match 28.3%; Score 157; DB 1; Length 4548;
Best Local Similarity 34.2%; Pred. No. 3.7e-09;
Matches 40; Conservative 9; Mismatches 36; Indels 32; Gaps 6;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSGATVLQOYTHAHRSDAL---QLGLGKHNYCRNP 59
DB 3896 CYRGDGSYRGTLSTTTIGTKCQSWSS-----MTPWHRIPLYPNAGLTR-NYCRNP 3949

QY 60 NRRPWCYVQVGLKPLV--QECMVHDC-----ADGKPSPPPE 96
DB 3950 AEIRPWCYT---MDPSVRWEYCNLTRCPVTSESSVLTPTVPVPESTAPSEQAPPEK 4003

RESULT 24
E61545
Plasmin (EC 3.4.21.7) precursor - dog (fragments)
N;Alternate names: plasminogen
C;Species: Canis lupus familiaris (dog)
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C;Accession: E61545
R;Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
Article: Structural aspects of the plasminogen of various species.
A;Reference number: A61545; MUID:89005015; PMID:3168975
A;Accession: E61545

A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-120 <SCH>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F;37-114/Domain: kringle homology <KR4>

Query Match 28.2%; Score 156; DB 2; Length 120;
Best Local Similarity 39.6%; Pred. No. 1.5e-09;
Matches 36; Conservative 7; Mismatches 30; Indels 18; Gaps 5;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSGATVLQOYTHAHRSDAL---QLGLGKHNYCRN 57
DB 37 CYHNGGSYRGTLSTTTIGTKCQSWSSMT-----PRHEKTEHPPEAGL-TMNYCRN 88

QY 58 PDNRRPWCYVQVGLKPLV--QECMVHDCAD 86
DB 89 PDADKSPWCYT---TDPSVRWEFCNLRKCLD 116

RESULT 25
S45281
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)
N;Alternate names: Hageman factor (activated)
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C;Accession: S45281; A61329
R;Shibuya, Y.; Semba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A;Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): comparison with the structure of human factor XII (Hageman factor).
A;Reference number: S45281; MUID:94242782; PMID:8186251
A;Accession: S45281
A;Molecule type: mRNA
A;Residues: 1-593 <SHI>
A;Cross-references: GB:S70164
A;Note: The authors translated the codon GAG for residue 23 as Val, GAG for residue 70 as Pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 as Gln, and ATC for residue 505 as Leu
R;Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A;Title: Isolation and characterization of bovine factor XII (Hageman factor).
A;Reference number: A61329; MUID:77182112; PMID:861210
A;Accession: A61329
A;Molecule type: protein
A;Residues: 10-16, 'X', 18-19;525-550 <FUJ>
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma;
F;37-78/Domain: fibronectin type II repeat homology <IF2>
F;88-120/Domain: EGF homology <EGF>
F;125-160/Domain: fibronectin type I repeat homology <FB1>
F;207-287/Domain: kringle homology <KR>
F;350-587/Domain: trypsin homology <TRY>
F;541/Active site: Ser #status predicted

Query Match 28.1%; Score 155.5; DB 2; Length 593;
Best Local Similarity 38.4%; Pred. No. 7.7e-09;
Matches 33; Conservative 9; Mismatches 3; Indels 7; Gaps 3;

QY 2 TCYE--GNGHFYRGKASTDTMGRCPLPWSGATVLQOYTHAHRSDALQGLGKHNYCRNP 58
DB 206 SCYDDRDGSLSYRGWAGTTLTGAPCQSWAS----EATYWNVTAEQVLNGLGDHAFCRNP 261

QY 59 DNRPRPWCYVQVGLKPLVQECMVHDC 84
DB 262 DNDTEPCWCFWKGDRLSNYCRLLAPC 287

RESULT 26
A60140
Plasmin (EC 3.4.21.7) precursor - chicken (fragment)
N;Alternate names: plasminogen
C;Species: Gallus gallus (chicken)
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 16-Jul-1999
A;Accession: E61545


```
QY 60 NRREPWCYVQVGLKPLV--QECMVHDCADGKK--PSSP 93
Db 438 ADKSPWCYT---TDPVRWFEFNLKCKSETPEQVPAAP 472

RESULT 29
C61545
Plasmin (EC 3.4.21.7) precursor - goat (fragments)
N/Alternate names: plasminogen
C/Species: Capra aegagrus hircus (domestic goat)
C/Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C/Accession: C61545
R/Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A/Title: Structural aspects of the plasminogen of various species.
A/Reference number: A61545; MUID:89005015; PMID:168975
A/Accession: C61545
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-123 <SCH>
C/Superfamily: plasmin; kringie homology; plasminogen-related protein precursor homology
C/Keywords: hydrolase; serine proteinase
F/34-118/Domain: kringie homology <KR4>

Query Match 27.1%; Score 150; DB 2; Length 123;
Best Local Similarity 37.5%; Pred. No. 6.7e-09;
Matches 33; Conservative 12; Mismatches 31; Indels 12; Gaps 4;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATV--LQOTYHAHRSDALQLGLGKKNYCRNPD 60
Db 41 CYHNGQSTRGTSSTTTGKQCSWSMIPHRHQKTPESYNAGLTM-----NYCRNPD 95

QY 61 RREPWCYVQVGLKPLV--QECMVHDCAD 86
Db 96 DKSEWCYT---TDPVRWFEFNLKCKSE 120

RESULT 30
A40522
plasmin (EC 3.4.21.7) precursor - rat (fragment)
C/Species: Rattus norvegicus (Norway rat)
C/Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
C/Accession: A40522
R/Kanalias J.J.; Makler, S.P.
J. Biol. Chem. 266, 10825-10829, 1991
A/Title: Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor
A/Reference number: A40522; MUID:91250378; PMID:1645711
A/Accession: A40522
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-169 <KAN>
A/Cross-references: GB:M62832; NID:G206215; PIDN:AAA1884.1; PID:G554488
A/Note: the authors translated the codon TCT for residue 76 as Ala
C/Superfamily: plasmin; kringie homology; plasminogen-related protein precursor homology
C/Keywords: fibrinolysis; kringie homology; hydrolase; kringie; serine proteinase
F/34-112/Domain: kringie homology <KRG>
F/34-112, 55-95, 83-107/Disulfide bonds: #status predicted

Query Match 27.1%; Score 150; DB 2; Length 169;
Best Local Similarity 34.3%; Pred. No. 9.1e-09;
Matches 37; Conservative 14; Mismatches 31; Indels 26; Gaps 7;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLOOTYHAHRSDALQL---GLGHKNYCRNPD 59
Db 34 CYOENGKSYRGTSSTTTGKQCSW-----VSWTPHSHSKTANFPDSDL-ENNYCRNPD 87

QY 60 NRRPWCYVQVGLKPLV--QECMVHDCAD-----GKKPSP 93
Db 88 NDORGWPCFT---TDPVRWFEFNLKRCSETGGVAESAIVQVPSAP 132

RESULT 31
```

```
B30848
Plasmin (EC 3.4.21.7) precursor - rhesus macaque
C/Species: Macaca mulatta (rhesus macaque)
C/Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 22-Jun-1999
C/Accession: B32869; B30848
R/Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A/Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A/Reference number: A32869; MUID:89174660; PMID:2925643
A/Accession: B32869
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-810 <TOM>
A/Cross-references: GB:J04697; NID:G342272; PIDN:AAA36901.1; PID:G342273
C/Superfamily: plasmin; kringie homology; plasminogen-related protein precursor homology
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringie; serine proteinase
F/1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F/1-9/Domain: signal sequence #status predicted <SIG>
F/103-181/Domain: kringie homology <KR1>
F/185-262/Domain: kringie homology <KR2>
F/275-352/Domain: kringie homology <KR3>
F/377-454/Domain: kringie homology <KR4>
F/481-560/Domain: kringie homology <KR5>
F/581-803/Domain: trypsin homology <TRY>
F/49-73, 53-61, 103-181, 124-164, 152-176, 185-262, 188-316, 206-245, 234-257, 275-352, 296-335, 3
bonds: #status predicted
F/622, 665, 760/Active site: His, Asp, Ser #status predicted

Query Match 26.9%; Score 149; DB 2; Length 810;
Best Local Similarity 36.4%; Pred. No. 5.2e-08;
Matches 36; Conservative 12; Mismatches 35; Indels 16; Gaps 6;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLOOTYHAHR---SDALQLGLGKKNYCRNPD 59
Db 377 CYHGDGQSYRGTSSTTTGKQCSWS-----MTPHWEKTPENFENAGL-TWNYCRNPD 430

QY 60 NRREPWCYVQVGLKPLV--QECMVHDCAD--DGKKPSPSP 94
Db 431 ADKGPWCFT---TDPVRWFEFNLKCKSETGSGVAAPPP 466

RESULT 32
A47136
macrophage-stimulating protein 1 precursor - human
C/Species: Homo sapiens (man)
C/Date: 03-May-1994 #sequence_revision 14-Nov-1997 #text_change 18-Jun-1999
C/Accession: A40331; B40331; A47136; A61395
R/Han, S.; Stuart, L.A.; Degen, S.J.F.
Biochemistry 30, 9768-9780, 1991
A/Title: Characterization of the DNFI5S2 locus on human chromosome 3: identification of
A/Reference number: A40331; MUID:92002016; PMID:1655021
A/Accession: A40331
A/Molecule type: DNA
A/Residues: 1-711 <HA1>
A/Cross-references: GB:M74179
A/Accession: B40331
A/Molecule type: mRNA
A/Residues: 1-711 <HA2>
A/Cross-references: GB:M74178; NID:G183976; PIDN:AAA50165.1; PID:G183977
R/Yoshimura, T.; Yuhki, N.; Wang, M.H.; Skeel, A.; Leonard, E.J.
J. Biol. Chem. 268, 15461-15468, 1993
A/Title: Cloning, sequencing, and expression of human macrophage stimulating protein (M
A/Reference number: A47136; MUID:93340141; PMID:8393443
A/Accession: A47136
A/Molecule type: mRNA
A/Residues: 1-12, 'C', 14-622, 'F', 624-711 <YOS>
A/Cross-references: GB:L11924; NID:G938037; PIDN:AAA59872.1; PID:G938038
A/Note: authors translated the codon TTT for residue 623 as Ieu; parts of this sequence
R/Skeel, A.; Yoshimura, T.; Showalter, S.D.; Tanaka, S.; Appella, E.; Leonard, E.J.
J. Exp. Med. 173, 1227-1234, 1991
A/Title: Macrophage stimulating protein: purification, partial amino acid sequence, and
A/Reference number: A61395; MUID:91217635; PMID:1827141
A/Accession: A61395
```

```

RESULT 34
PLPG
plasmin (EC 3.4.21.7) precursor - pig (fragment)
N:Alternate names: plasminogen
N:Contains: miniplasminogen
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 07-Sep-1990 #sequence revision 01-Nov-1996 #text_change 18-Jul-1997
C:Accession: S03733; S03737; A25834
R:Schaller, J.; Marti, T.; Roessleat, S.J.; Kaempfer, U.; Rickli, E.E.
Fibrinolysis 1, 91-102, 1987
A:Title: Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the ca
A:Reference number: S03733
A:Accession: S03733
A:Molecule type: protein
A:Residues: 1-560 <SCH>
R:Brundish, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;
Eur. J. Biochem. 114, 465-470, 1981
A:Title: Comparison of the primary structure of the N-terminal CNBR fragments of human, i
A:Reference number: S03735; MUID:81212097; PMID:7238497
A:Accession: S03737
A:Molecule type: protein
A:Residues: 1-57 <BRU>
R:Marti, T.; Schaller, J.; Rickli, E.E.
Eur. J. Biochem. 149, 279-285, 1985
A:Title: Determination of the complete amino-acid sequence of porcine miniplasminogen.
A:Reference number: A25834; MUID:85203907; PMID:3846533
A:Accession: A25834
A:Molecule type: protein
A:Residues: 450-790 <NAR>
C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen act
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma; serine prote
F:1-790/Product: plasminogen #status predicted <PRO>
F:1-77/Domain: plasminogen-related protein precursor homology (fragment) <PLPH>
F:78-560/Product: plasmin chain A #status predicted <APT>
F:84-162/Domain: kringle homology <KR1>
F:166-243/Domain: kringle homology <KR2>
F:256-333/Domain: kringle homology <KR3>
F:358-435/Domain: kringle homology <KR4>
F:450-790/Product: miniplasminogen #status experimental <MIN>
F:461-540/Domain: kringle homology <KR5>
F:561-783/Domain: trypsin homology <TRY>
F:30-54, 34-42, 84-162, 105-145, 133-157, 166-243, 169-297, 187-226, 215-238, 256-333, 277-316, 305
bonds: #status predicted
F:502,645,740/Active site: His, Asp, Ser #status predicted

Query Match 26.7%; Score 148; DB 1; Length 790;
Best Local Similarity 38.2%; Pred. No. 6.4e-08;
Matches 34; Conservative 8; Mismatches 33; Indels 14; Gaps 5;

QY 3 CYEGNGHYRGKASTDTWGRCLPWN SATVLOOTYHAHR---SDALQILGKHNYCRNPD 59
DB 358 CYRGNGESYRGKSTITITGRKQSW-----VSMTPHRHKTGPNFPNAGL-TMNYCRNPD 411
QY 60 NRRRRPWCYVQGLKPLV--QECMVHDCAD 86
DB 412 ADKSPWCYT---TDRVRWEYCNLKKQSE 437

RESULT 35
A40332
macrophage-stimulating protein 1 precursor - mouse
N:Alternate names: hepatocyte growth factor-like protein
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 18-Jun-1999
C:Accession: A40332; B40332
R:Degen, S.J.F.; Stuart, L.A.; Han, S.; Jamison, C.S.
Biochemistry 30, 9781-9791, 1991

```

A;Title: Characterization of the mouse cDNA and gene coding for a hepatocyte growth factor
A;Reference number: A40332; MUID:92002017; PMID:1832957
A;Accession: A40332
A;Molecule type: DNA
A;Residues: 1-716 <DEG>
A;Cross-references: GB:M74180; NID:gl93831; PIDN:AAA50166.1; PID:gl93832
A;Accession: B40332
A;Molecule type: mRNA
A;Residues: 1-18, 'P', 20-716 <DEG2>
A;Cross-references: GB:M74181; NID:gl93833; PIDN:AAA50167.1; PID:gl93834
C;Genetics:
A;Introns: 18/1; 67/2; 105/1; 143/2; 189/1; 229/2; 269/1; 334/2; 378/1; 412/2; 458/1; 47
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: duplication; glycoprotein; growth factor; kringle
F;1-31/Domain: signal sequence #status predicted <SIG>
F;19-488,489-716/Product: macrophage-stimulating protein 1 #status experimental <MAT>
F;110-483/Domain: alpha chain #status experimental <ACH>
F;110-186/Domain: kringle homology <XR1>
F;191-268/Domain: kringle homology <XR2>
F;292-370/Domain: kringle homology <XR3>
F;379-457/Domain: kringle homology <XR4>
F;484-711/Domain: beta chain #status experimental <BCH>
F;489-709/Domain: trypsin homology <TRY>
F;72,173,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 26.6%; Score 147.5; DB 1; Length 716;
Best Local Similarity 36.5%; Pred. No. 6.6e-08;
Matches 38; Conservative 11; Mismatches 34; Indels 21; Gaps 7;
CY 3 CYEGNGHFYRGKASTDTMGRPCLPWN SATVLOQTYHAHR-----SDALQLGLGKGYNCRN 57
DB 379 CYHSGGEYRGVSVKTKRGVQCQHWSET-----PHKPPQTFPSAPQAGL-EANFCRN 430
CY 58 PD-NRRPWCYVQVGLKP--LVQECMVHDCADGKKPS--SPPEE 96
DB 431 PDGDSHGFWCYT---LDPDILFYCALQRCDDQDPESILDPDPQ 471
neurotrophic receptor ror2 precursor - human
A;Title: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 19-May-2000
A;Accession: B45082
R;Maslowski, P.; Carroll, R.D.
J. Biol. Chem. 267, 26181-26190, 1992
A;Title: A novel family of cell surface receptors with tyrosine kinase-like domain.
A;Reference number: A45082; MUID:93100347; PMID:1334494
A;Accession: B45082
A;Molecule type: mRNA
A;Residues: 1-943 <MAS>
A;Cross-references: GB:M97639; NID:gl337466; PIDN:AAA60276.1; PID:gl337467
A;Note: sequence extracted from NCBI backbone (NCBIP:120918)
C;Genetics:
A;Gene: GDB:NTRKR2
A;Cross-references: GDB:1136454
A;Map position: 6p21-6p21
C;Superfamily: neurotrophic receptor ror; immunoglobulin homology; kringle homology; pro
C;Keywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyros
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-937/Product: neurotrophic receptor ror2 #status predicted <MAT>
F;76-137/Domain: immunoglobulin homology <INM>
F;316-394/Domain: kringle homology <KRG>
F;412-428/Domain: transmembrane #status predicted <TMN>
F;471-753/Domain: protein kinase homology <KIN>
F;479-487/Region: protein kinase ATP-binding motif
F;70,188,318/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 26.6%; Score 147.5; DB 2; Length 943;
Best Local Similarity 39.3%; Pred. No. 8.6e-08;
Matches 33; Conservative 10; Mismatches 36; Indels 5; Gaps 4;

CY 3 CYEGNGHFYRGKASTDTMGRPCLPWN SATVLOQTYHAHRSDALQLGLGKGYNCRNPNRR 62
DB 316 CYNGSGVYRGTAATTKSGHCCQPW--ALQPHSHLSSTDFPELG--GGHAYCENFGQOM 372
CY 63 R-PWCYVQVGLKPLVQECMVHDC 85
DB 373 EGPWCFTQ-NKNVRMELCDVPSCS 395

RESULT 37

PLHU
A;Title: precursor [validated] - human
N;Alternate names: plasminogen precursor [misnomer]
N;Contains: angiotatin; microplasmin; plasminogen
C;Species: Homo sapiens (man)
C;Date: 24-Apr-1994 #sequence_revision 02-Dec-1994 #text_change 15-Sep-2000
C;Accession: A35229; I52242; K26646; I62738; I84609; S03735; A00929; A04627; A04625; A0
R;Petersen, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
J. Biol. Chem. 265, 6104-6111, 1990
A;Title: Characterization of the gene for human plasminogen, a key proenzyme in the fib
A;Reference number: A35229; MUID:90202879; PMID:2318848
A;Accession: A35229
A;Molecule type: DNA
A;Residues: 1-810 <PET>
A;Cross-references: GB:J05286; GB:M34276; NID:gl90064; PIDN:AAA60113.1; PID:gl90064
A;Experimental source: leukocyte; lung fibroblast
R;Malgaroli, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.; T
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
A;Title: Definition of the transcription initiation site of human plasminogen gene in 1
A;Reference number: I52242; MUID:91097523; PMID:2268308
A;Accession: I52242
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-16 <MAL1>
A;Cross-references: GB:M62890; NID:gl90092; PIDN:AAA36454.1; PID:gl90092
R;Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FEBS Lett. 213, 254-260, 1987
A;Title: Molecular cloning and characterization of a full-length cDNA clone for human p
A;Reference number: A26646; MUID:87162490; PMID:3030813
A;Accession: A26646
A;Molecule type: mRNA
A;Residues: 1-471, 'D', 473-810 <FOR>
A;Cross-references: GB:X05199; NID:gl35530; PIDN:CAA28831.1; PID:gl35531
A;Experimental source: liver
R;Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A;Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A;Reference number: I45961; MUID:85023311; PMID:6148961
A;Accession: I62738
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 292-471, 'D', 473-810 <MAL2>
A;Cross-references: GB:K02922; NID:gl90112; PIDN:AAA60124.1; PID:gl907031
A;Accession: I84609
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 367-419 <MAL3>
A;Cross-references: GB:K02921; NID:gl90110; PIDN:AAA60123.1; PID:gl90111
R;Brühnholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.
Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBR fragments of human,
A;Reference number: S03735; MUID:81212097; PMID:7238497
A;Accession: S03735
A;Molecule type: protein
A;Residues: 20-71, 'E', 73-76 <BRU>
R;Søstrup-Jensen, L.; Petersen, T.E.; Magnusson, S.
submitted to the Atlas, July 1977
A;Reference number: A00929
A;Accession: A00929
A;Molecule type: Protein
A;Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>
R;Wiman, B.

A:Reference number: A65245; PDB:1CEB
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R:Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.
Biochemistry 30, 10576-10586, 1991
A:A>Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 Å
A:Reference number: A58819; MUID:92031502; PMID:1657148
A:Contents: annotation
R:Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.
Biochemistry 30, 10589-10594, 1991
A:A>Title: The refined structure of the epsilon-aminocaproic acid complex of human plasminogen
A:Reference number: A58818; MUID:92031503; PMID:1657149
A:Contents: annotation
R:Ride Vos, A.M.; Ullrich, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, M.L.
Biochemistry 31, 270-279, 1992
A:A>Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at 2.4 Å
A:Reference number: A39483; MUID:92118803; PMID:1310033
A:Contents: annotation; X-ray crystallography, 2.4 angstroms
R:Stec, B.; Tester, M.M.; Whitlow, M.; Yamano, A.
submitted to the Brookhaven Protein Data Bank, June 1995
A:Reference number: A65980; PDB:1KRN
A:Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454
R:Rejante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A65804; PDB:1HPK
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181
R:Rejante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 927-937, 1994
A:A>Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
A:Reference number: S43645; MUID:94237157; PMID:8181475
A:Contents: annotation; conformation by (1)H-NMR, residues 96-184
R:Rejante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 939-949, 1994
A:A>Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasminogen
A:Reference number: A58817; MUID:94237158; PMID:8181476
A:Contents: annotation; conformation by (1)H-NMR
C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many other fluids.
C:Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:UKHU a d PIR:FGUHB).
C:Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITHUA2) immediately after release from fibrin clots.
C:Comment: Microplasmin is formed by autolytic cleavage of plasmin under artificial conditions.
C:Comment: Streptolysin 1 (see PIR:KCHUS1) acts on plasminogen to produce angiostatin. Toxicity solid tumors.
C:Genetics:
A:Gene: GDB:PLG
A:Cross-references: GDB:119498; OMIM:173350
A:Map position: 6Q26-Q27
A:Introns: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 529/1;
C:function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in various tissues; also activates the urokinase-type plasminogen activator.
A:Pathway: fibrinolysis
A:Superfamily: plasmin
C:Keywords: angiotensin inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydrolysis
F:1-96/Domain: plasminogen-related protein precursor homology
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-810/Product: plasminogen #status experimental <PRO>
F:20-96/Domain: activation peptide #status experimental <APT>
F:79-466/Product: angiotensin #status experimental <AS>
F:97-580,581-810/Product: plasmin #status experimental <MAT>
F:97-580/Domain: plasmin chain A #status experimental <CHA>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:550-580,581-810/Product: microplasmin #status experimental <MMT>

Query Match 26.4%; Score 146.5; DB 1; Length 810;

Query Match 26.4%; Score 146.5; DB 1; Length 810;

Best Local Similarity 37.4%; Pred. No. 9.5e-08;
Matches 37; Conservative 13; Mismatches 32; Indels 17; Gaps 7;
QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPD 59
Db 377 CYHGDSQYRGVTSITTTCKKQSHSS-----WTPRHQKTPENYFNAGL-TMNYCRNPD 430
QY 60 NRRRPWCYVQVGLKPLV--OECVHDCADGKKPS--SPP 94
Db 431 ADKGWPCFT---TDPVRWEYCNLKCS-GTEASVWAPP 465
RESULT 38
A45082
neurotrophic receptor ror1 precursor - human
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 19-May-2000
C;Accession: A45082
R;Maslakowski, P.; Carroll, R.D.
J. Biol. Chem. 267, 26181-26190, 1992
A;Title: A novel family of cell surface receptors with tyrosine kinase-like domain.
A;Reference number: A45082; MUID:93100347; PMID:1334494
A;Accession: A45082
A;Molecule type: mRNA
A;Residues: 1-937 <MAS>
A;Cross-references: GB:M97675; NID:9337464; PIDN:AAA60275-1; PID:9337465
A;Note: sequence extracted from NCBI backbone (NCBIP:120916)
C;Genetics:
A;Gene: GDB:NTRK1
A;Cross-references: GDB:136453
A;Map position: 6p21-6p21
C;Superfamily: neurotrophic receptor ror; immunoglobulin homology; kringle homology; pro
C;Keywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyros
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-937/Product: neurotrophic receptor ror1 #status predicted <MAT>
F;72-133/Domain: immunoglobulin homology <IMM>
F;313-391/Domain: kringle homology <XRG>
F;404-425/Domain: transmembrane #status predicted <TM1>
F;471-753/Domain: protein kinase homology <KIN>
F;47,66,184,315/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 26.4%; Score 146; DB 2; Length 937;
Best Local Similarity 44.1%; Pred. No. 1.2e-07;
Matches 30; Conservative 5; Mismatches 25; Indels 8; Gaps 3;
QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLG--LGKHNYCRNPD 60
Db 313 CYNSTGVYRGTVSVTKSGRCQCPWNS-----QYPHTHTFTALRPPELNGHSYCNPN 367
QY 61 RRR-PMCY 67
Db 368 QKEAPWCF 375
RESULT 39
T18840
hypothetical protein CO1G6.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T18840
R;Barks, M.
submitted to the EMBL Data Library, August 1994
A;Reference number: Z19029
A;Accession: T18840
A;Status: preliminary; translated from GB/EMBL/DDJ
A;Molecule type: DNA
A;Residues: 1-806 <WIL>
A;Cross-references: EMBL:Z35595; PIDN:CAA84639-1; GSPDB:GN00020; CESP:CO1G6.8
A;Experimental source: clone CO1G6
C;Genetics:
A;Gene: CESP:CO1G6.8
A;Map position: 2

A;Introns: 36/3; 170/3; 217/3; 636/3; 760/1
Query Match 26.2%; Score 145; DB 2; Length 806;
Best Local Similarity 35.2%; Pred. No. 1.4e-07;
Matches 32; Conservative 12; Mismatches 35; Indels 12; Gaps 5;
QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPD 62
Db 232 CYNSTGVYRGTVSVTKSGRCQCPWIDST--SRDFNVHRPEL---MNSKNYCRNPGKK 286
QY 63 -RPMCYVQVGLKPLVQE--CMVHDCADGKKP 90
Db 287 SRPCY----SKPMGQEYCDVQCPSDMYP 313
RESULT 40
PLMS
plasmin (EC 3.4.21.7) precursor - mouse
N;Contains: angiotatin; plasminogen
C;Species: Mus musculus (house mouse)
C;Date: 20-Sep-1991 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C;Accession: A38514; S48202; S48203
R;Degen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.
Genomics 8, 49-61, 1990
A;Title: Characterization of the cDNA coding for mouse plasminogen and localization of
A;Reference number: A38514; MUID:91184812; PMID:2081600
A;Accession: A38514
A;Molecule type: mRNA
A;Residues: 1-812 <DEG>
A;Cross-references: GB:J04766; NID:G200402; PIDN:AAA50168-1; PID:G200403
R;Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A;Title: Characterization of the murine plasma fibrinolytic system.
A;Reference number: S48202; MUID:95010076; PMID:7523120
A;Accession: S48202
A;Molecule type: protein
A;Residues: 20-25 <LIJ>
A;Accession: S48203
A;Molecule type: protein
A;Residues: 22-27 <LIJ>
C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many o
C;Comment: Plasminogen is converted into plasmin by plasminogen activators, both plasm
e inhibitor, the activation involves also removal of the activation peptide.
C;Comment: Stremelysin 1 (see PIC:KCMSS1) acts on plasminogen to produce angiotatin. I
eul in treating solid tumors.
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a v
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen ac
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolog
C;Keywords: angiotatin inhibitor; blood; duplication; fibrinolysis; glycoprotein; hyd
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-812/Product: plasminogen #status predicted <PRO>
F;20-96/Domain: activation peptide #status predicted <APT>
F;79-466/Product: angiotatin #status predicted <AST>
F;97-581,582-812/Product: plasmin #status predicted <MAT>
F;97-581/Domain: chain A #status predicted <ACH>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;377-454/Domain: kringle homology <KR4>
F;481-560/Domain: kringle homology <KR5>
F;582-812/Domain: chain B #status predicted <BCH>
F;582-805/Domain: trypsin homology <TRY>
F;49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,3
bonds: #status predicted
F;78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted
F;136,308/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted
F;581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental
F;624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 26.1%; Score 144.5; DB 1; Length 812;
Best local similarity 32.7%; Pred. No. 1.6e-07;
Matches 33; Conservative 8; Mismatches 47; Indels 13; Gaps 4;

QY 3 CYEGNGHFYRGKASTDTMGPCPLPWSATVLOQTYHAHRSDALQLGLG--KKNYCRNPDN 60
DB 275 CLKRGENYRGTVSVTVSGTKCQWMS-----EQTPHRHNRTPENFPCKNLENYCRNPDG 329

QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKXP-----SSPPEE 96
DB 330 ETAPWCYT-TDSQLRWYCEIPSCSSASPDQSDSDSVPPPEE 369

RESULT 41
T18518
apolipoprotein(a) - western European hedgehog (fragment)
C/Species: Erinaceus europaeus (western European hedgehog)
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
C/Accession: T18518
R/Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong
J. Biol. Chem. 270, 24004-24009, 1995
A/Title: The recurring evolution of lp(a): Insights from cloning of hedgehog apolipoprotein
A/Reference number: I46259; MUID:96025778; PMID:7592597
A/Accession: T18518
A/Status: preliminary; translated from GB/EMBL/DBDB
A/Molecule type: mRNA
A/Residues: 1-2869 <LAW>
A/Cross-references: EMBL:U33170; NID:g1046358; PID:g1046359; PIDN:AAC48522.1
A/Experimental source: liver
C/Comment: The lipoprotein lp(a), a major inherited risk factor for atherosclerosis, contains apolipoprotein(a).

Query Match 25.8%; Score 143; DB 2; Length 2869;
Best local similarity 35.1%; Pred. No. 7.5e-07;
Matches 33; Conservative 8; Mismatches 45; Indels 8; Gaps 4;

QY 3 CYEGNGHFYRGKASTDTMGPCPLPWSATVLOQTYHAHRSDALQLGLGKKNYCRNPDRR 62
DB 2591 CLENGENYQGNMAITVSGPCQWRKQTQPHRHEVTPENYPSKNL-FG--NYCRNPDEGI 2647

QY 63 RPYCYVQVGLKPLVQECMVHDCADGKKSPSPPEE 96
DB 2648 APWCYT-TNSAVRWYCSITFC-----ESSSPPTTE 2676

RESULT 42
A61545
plasmin (EC 3.4.21.7) precursor - horse (fragments)
N/Alternate names: plasminogen
N/Contains: miniplasminogen
C/Species: Equus caballus (domestic horse)
C/Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
C/Accession: A61545; S17527
R/Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A/Title: Structural aspects of the plasminogen of various species.
A/Reference number: A61545; MUID:9205015; PMID:3168975
A/Accession: A61545
A/Molecule type: protein
A/Residues: 1-33,34-117 <SCH>
R/Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 4, 69-74, 1991
A/Title: Complete amino acid sequence of equine miniplasminogen.
A/Reference number: S17527; MUID:92052077; PMID:1946332
A/Accession: S17527
A/Molecule type: protein
A/Residues: 118-455 <SC2>
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C/Keywords: fibrinolysis; Glycoprotein; hydrolase; kringe; plasma; serine proteinase; z
F/1-33,34-117,118-455/Product: plasminogen (fragments) #status experimental <PRO>
F/1-33/Domain: activation peptide (fragment) #status experimental <APT>
F/34-117,118-225,226-455/Product: plasmin (fragments) #status experimental <NAT>

```

F;37-114/Domain: kringle homology <KR4>
F;118-455/Product: miniplasminogen #status experimental <MIN>
F;126-205/Domain: kringle homology <KR5>
F;226-455/Domain: plasmin chain B #status experimental <BCH>
F;226-448/Domain: trypsin homology <TRY>
F;267,310,405/Active site: His, Asp, Ser #status predicted

Query Match          25.5%; Score 141.5; DB 2; Length 455;
Best Local Similarity 34.4%; Pred. No.1.9e-07;
Matches 33; Conservative 15; Mismatches 35; Indels 13; Gaps 5;

QY 3 CYEENGHFYRGKASTDTMGRCPLPNWNSATV--LQQTYYHAHRSDALQLGLGKHNYCRNPDN 60
Db 37 CYQDKGESYRGTSIIIVTKGKQCSWSTPHWHQHTPEKYPNADLTM-----NYCRNPDG 91
QY 61 RRRPWCYVQVGLKPLV--QECMVHDCADG-KKPSSP 93
Db 92 DKGPWCYT---TDPSVWEFCNLKCKSETVQEPSEP 124

neurotrophic receptor ror precursor - fruit fly (Drosophila melanogaster)
N:Alternate names: trk-related receptor
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Drosophila melanogaster
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 17-Nov-2000
C:Accession: A48289
A:Wilson, C.; Goberdhan, D.C.I.; Steller, H.
Proc. Natl. Acad. Sci. U.S.A. 90, 7109-7113, 1993
A:Title: Ror, a potential neurotrophic receptor gene, encodes a Drosophila homolog of t
A:Reference number: A48289; MUID:93348222; PMID:8394009
A:Accession: A48289
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-685 <MIL>
A:Cross-references: GB:L20297; NID:5348103; PIDN:AAA28860.1; PID:G348104
C:Genetics:
A:Gene: FlyBase:bsk
A:Cross-references: FlyBase:FBgn010407
C:Superfamily: Drosophila neurotrophic receptor ror; kringle homology; protein kinase hom
C:Keywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyros
F;237-310/Domain: kringle homology <KRG>
F;334-338/Domain: transmembrane #status predicted <TM1>
F;408-677/Domain: protein kinase homology <KIN>
F;416-424/Region: protein kinase ATP-binding motif
F;45,63,123,144,250/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          25.1%; Score 139; DB 1; Length 685;
Best Local Similarity 34.5%; Pred. No.5.2e-07;
Matches 30; Conservative 19; Mismatches 28; Indels 10; Gaps 5;

QY 1 KTYEENGHFYRGKASTDTMGRCPLPNWNSATVLLQQTYYHAHRSDALQLGLGKHNYCRNPDN 60
Db 235 ENCYWEDGSGTYRGVANVSAGKPCLRW---SWLMKEI-----SDFPEL-IGQ-NYCRNPGS 285
QY 61 -RRRPWCYVQVGLKPLVQECMVHDCAD 86
Db 286 VENSFWCFVDSSRRIELCDIPKCAD 312

macrophage-stimulating protein 1 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Jun-2000
C:Accession: JC5061
F;Oshiro, K.; Iwama, A.; Matsuno, K.; Ezaki, T.; Sakamoto, O.; Hamaguchi, I.; Takasu, N.
Biochem. Biophys. Res. Commun. 227, 273-280, 1996
A:Title: Molecular cloning of rat macrophage-stimulating protein and its involvement in
A:Reference number: JC5061; MUID:97011126; PMID:8858136
A:Accession: JC5061
A:Molecule type: mRNA

```

R.Coffer, A.; Fellows, J.; Young, S.; Pappin, D.; Rahman, D.
Biochem. J. 278, 35-41, 1991
A:Title: Purification and characterization of biologically active scatter factor from
A:Reference number: S17173; MUID:91354223; PMID:1831975
A:Accession: S17173
A:Molecule type: protein
A:Residues: 496-517, 'T', 519 <COF>
R:Gherardi, E.; Stoker, M.
Nature 346, 228, 1990
A:Title: Hepatocytes and scatter factor.
A:Reference number: S10966; MUID:90326152; PMID:2142751
A:Accession: S10966
A:Status: preliminary
A:Molecule type: protein
A:Residues: 496-507, 'X', 509-512, 'L', 514-516, 'X', 518-519 <NAT>
R:Plaschke-Schlutter, A.; Behrens, J.; Gherardi, E.; Birchmeier, W.
J. Biol. Chem. 270, 830-836, 1995
A:Title: Characterization of the scatter factor/hepatocyte growth factor gene promoter.
A:Reference number: I48758; MUID:95122532; PMID:7822318
A:Accession: I48758
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-30 <RES>
A:Cross-references: EMBL:X81630; NID:G673451; PIDN:CRA57286.1; PID:G673452
A:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Function:
A:Description: stimulates mitosis of hepatocytes and other cells
A:Note: does not have proteinase activity
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; py
F:1-32/Domain: signal sequence #status predicted <SIG>
F:56-495/Domain: product: hepatocyte growth factor #status predicted <MAT>
F:56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F:129-207/Domain: kringle homology <K1>
F:212-289/Domain: kringle homology <K2>
F:306-384/Domain: kringle homology <KR3>
F:392-470/Domain: kringle homology <KR4>
F:496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
F:496-719/Domain: trypsin homology <TRY>
F:33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:295,403,569,656/Binding site: carboxylate (Asn) (covalent) #status predicted
F:488-607/Disulfide bonds: #status predicted

Query Match 24.8%; Score 137.5; DB 1; Length 728;
Best Local Similarity 32.0%; Pred. No. 7,9e-07;
Matches 31; Conservative 12; Mismatches 35; Indels 19; Gaps 4;

QY 3 CYEGNGHFYRGKASDTMGRPLPWNASATVLQQTTHAHRSDALQLGLG----KHNYCRNP 58
DB 306 CLQGGEGYRGTSNTWNGIPQRWDS-----QYFKHDDITPENFKCDLRENYCRNP 358
QY 59 DNRPRWCY-----VOVGLKPLVQECNV---HDCADG 87
DB 359 DGAESPWCFTTDPNIRVGYCSQIPKCDVSSGGDCYRG 395

RESULT 46
A35644
hepatocyte growth factor precursor - rat
N:Alternate names: hepatolentin A; scatter factor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 21-Jul-2000
C:Accession: A35644; S13211
R:Tashiro, K.; Hagiya, M.; Nishizawa, T.; Seki, T.; Shimonishi, M.; Shimizu, S.; Nakamu
Proc. Natl. Acad. Sci. U.S.A. 87, 3200-3204, 1990
A:Title: Deduced primary structure of rat hepatocyte growth factor and expression of th
A:Reference number: A35644; MUID:90222197; PMID:2139229
A:Accession: A35644
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-728 <RAS>
A:Cross-references: GB:D90102; GB:M32987; NID:g220766; PIDN:BAAL14133.1; PID:g220767
A:Note: the authors translated the codon GAG for residue 70 as Gln. GAG for residue 417

.....

R;Okajima, A.; Miyazawa, K.; Kitamura, N.
Eur. J. Biochem. 193, 375-381, 1990
A;Title: Primary structure of rat hepatocyte growth factor and induction of its mRNA during
A;Reference number: S13211; MUID:91031482; PMID:2146117
A;Accession: S13211
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-728 <OKA>
A;Cross-references: EMBL:X54400; NID:G56353; PIDN:CAA38266.1; PID:G4539554
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Function:
A;Description: stimulates mitosis of hepatocytes and other cells
A;Note: does not have proteinase activity
C;Superfamily: hepatocyte growth factor; kringlike homology; trypsin homology
C;Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringlike; pyH
F;1-32/Domain: signal sequence #status predicted <SIG>
F;56-495/Product: hepatocyte growth factor #status predicted <WAT>
F;56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F;129-207/Domain: kringlike homology <KR1>
F;212-289/Domain: kringlike homology <KR2>
F;306-384/Domain: kringlike homology <KR3>
F;392-470/Domain: kringlike homology <KR4>
F;496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
F;496-719/Domain: trypsin homology <TRY>
F;33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;295,403,569,656/Binding site: carboxylate (Asn) (covalent) #status predicted
F;488-607/Disulfide bonds: #status predicted

Query Match 24.5%; Score 135.5; DB 1; Length 728;
Best Local Similarity 32.0%; Pred. No. 1.3e-06;
Matches 31; Conservative 11; Mismatches 36; Indels 19; Gaps 4;
Qy 3 CYEGNGHYFGKASTDTMGPRCLPWSATVLQTYHAHRSDALQLGLG----KHNYCRNP 58
Db 306 CIRQGGGGRGTNTIWNIPQCRWDS-----QYPKHDTIPENFKCKDLRENYCRNP 358
Qy 59 DNRRRPCY-----VOVGLKPLVQECMV---HDCADG 87
Db 359 DGASPCWCTDPNIRVGYCSQIPKCDVSSQDCYRG 395

RESULT 47
151285
hepatocyte growth factor/scatter factor - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C;Accession: I51285
R;Streit, A.; Stern, C.D.; Thery, C.; Ireland, G.W.; Aparicio, S.; Sharpe, M.J.; Gherard
Development 121, 813-824, 1995
A;Title: A role for HGF/SF in neural induction and its expression in Hensen's node during
A;Reference number: I51285; MUID:95237013; PMID:7720585
A;Accession: I51285
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-411 <STR>
A;Cross-references: GB:S77480; NID:G998675; PID:G998676
C;Superfamily: hepatocyte growth factor; kringlike homology; trypsin homology
F;124-197/Domain: kringlike homology <KR3>
F;202-279/Domain: kringlike homology <KR2>
F;296-374/Domain: kringlike homology <KR3>

Query Match 24.1%; Score 133.5; DB 2; Length 411;
Best Local Similarity 32.0%; Pred. No. 1.2e-06;
Matches 32; Conservative 12; Mismatches 41; Indels 15; Gaps 4;
Qy 2 TCYEGNGHYFGKASTDTMGPRCLPWSATVLQTYHAHR--SDALQGLGKHNYCRNP 59
Db 295 TCICQGGGGRGTNTIWSGIQCRWDS-----QPPHONITPENFKCKDLRENYCRNP 349
Qy 60 NRRRPYCY-----VOVGLKPLVQECMV---HDCADGKPS 91
Db 350 GSSEPCWCTDPNIRIGYCSQIPKCDVSNEDQCYRGNGKS 389

RESULT 48

JH0579
hepatocyte growth factor precursor [validated] - human
N;Alternate names: hepatopietin A; scatter factor
C;Species: Homo sapiens (man)
C;Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 08-Dec-2000
C;Accession: JH0579; J00333; A41140; B36677; A36677; A33512; A39006; PH0114; A37796; S06;
R;Seki, T.; Hagiya, M.; Shimomishi, M.; Nakamura, T.; Shimizu, S.
Gene 102, 213-219, 1991
A;Title: Organization of the human hepatocyte growth factor-encoding gene.
A;Reference number: JH0579; MUID:91340155; PMID:1831432
A;Accession: JH0579
A;Molecule type: DNA
A;Residues: 1-728 <SEK>
A;Cross-references: DDBJ:D90318
A;Note: the authors translated the codon GAA for residue 662 as Gly
R;Seki, T.; Hagiya, M.; Shimomishi, M.; Nakamura, T.; Shimizu, S.
Submitted to JPIID, March 1991
A;Description: Organization of the human hepatocyte growth factor-encoding gene.
A;Reference number: J00333
A;Accession: J00333
A;Molecule type: DNA
A;Residues: 1-481, 'RT', 484-728 <SE2>
R;Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder, H.
Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991
A;Title: Evidence for the identity of human scatter factor and human hepatocyte growth f.
A;Reference number: A41140; MUID:91334393; PMID:1831266
A;Accession: A41140
A;Molecule type: mRNA
A;Residues: 1-728 <WEI>
A;Cross-references: GB:M73239; NID:G337935; PIDN:AAA64239.1; PID:G337936
R;Seki, T.; Ihara, I.; Sugimura, A.; Shimomishi, M.; Nishizawa, T.; Asami, O.; Hagiya, M.
Biochem. Biophys. Res. Commun. 172, 321-327, 1990
A;Title: Isolation and expression of cDNA for different forms of hepatocyte growth facto.
A;Reference number: A36677; MUID:91025062; PMID:2145836
A;Accession: B36677
A;Molecule type: mRNA
A;Residues: 1-728 <SE3>
A;Cross-references: GB:M60718; NID:G184031; PIDN:AAA52648.1; PID:G184032
A;Accession: A36677
A;Molecule type: mRNA
A;Residues: 1-161,167-728 <SE4>
A;Cross-references: EMBL:X16323
A;Experimental source: leukocyte
R;Miyazawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, N.; Arakaki, N.; Nakaya
Biochem. Biophys. Res. Commun. 163, 967-973, 1989
A;Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth fac
A;Reference number: A33512; MUID:89392017; PMID:2528952
A;Accession: A33512
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-728 <MIY>
A;Cross-references: GB:M29145; NID:G184041; PIDN:AAA52650.1; PID:G305846
R;Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.; Hir
Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991
A;Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepatocy
A;Reference number: A39006; MUID:91110540; PMID:1824873
A;Accession: A39006
A;Molecule type: mRNA
A;Residues: 1-161,167-728 <RUB>
A;Cross-references: GB:M55379
A;Experimental source: embryonic lung
R;Toshiyama, Y.; Arakaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayama, J
Biochem. Biophys. Res. Commun. 175, 660-667, 1991
A;Title: Identification of the N-terminal residue of the heavy chain of both native and
A;Reference number: PH0114; MUID:91207365; PMID:1826837
A;Accession: PH0114
A;Molecule type: protein
A;Residues: 32-43;53-58 <YOS>
A;Experimental source: plasma
R;Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier, W.
J. Cell Biol. 111, 2097-2108, 1990

Dd 305 CIOQGQEGVGVNTWTWNGIPQRWDN-----QYPHEHDMTPENFKCDLRNYCRNPDG 359

Qy 61 RRRPWCY-----VVQLGLPLVOEC-MVH--DCADG 87
||| : ||| : | | | | |
Dd 360 SESFPCFTTDPNIIRVGVCSPINCDSMHSQGDCVRG 394

RESULT 49

S33879
plasmin precursor - lamprey (fragments)
N;Alternate names: plasminogen
C;Species: Petromyzontidae gen. sp. (lamprey)
C;Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 07-Nov-1997
C;Accession: S33879
R;Afforter, M.; Schaller, J.; Rickli, E.E.
Protein Seq. Data Anal. 5, 207-211, 1993
A>Title: Isolation, characterization and partial amino acid sequence of lamprey plasmin
A;Reference number: S33879
A;Accession: S33879
A>Status: preliminary
A:Molecule type: protein
A;Residues: 1-1516-3435-44;45-59;60-76;77-111;111-138;139-158;159-176;179-216;217-236
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolog
F;81-146/Domain: kringle homology <KR3>

Query Match 23.6%; Score 131; DB 2; Length 336;
Best Local Similarity 32.9%; Pred. No. 1.9e-06;
Matches 28; Conservative 5; Mismatches 36; Indels 16; Gaps 2;

Qy 3 CYSGNHFYRGKASTDTMGPRCLPWNSATVLQOTYHAHRSDALQLGLGKNKYCNPNRR 62
: ||| : ||| : | | | | |
Dd 81 CVKGTGEGRGTAAITVSGKACAWASQT-----PGDVYSQGLSVSYCNPNPDGEK 131
: ||| : ||| : | | | | |

Qy 63 RPWCYYVQVGLKLPLVOECMVHDCCAG 87
||||| : ||| : | | | | |

Dd 132 LPWCYT-----TEYCWNPSCTGG 149
||||| : ||| : | | | | |

RESULT 50

I51283
hepatocyte growth factor precursor - clawed frog
N;Alternate names: hepatoin A; scatter factor
C;Species: Xenopus sp. (clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C;Accession: I51283
R;Nakamura, H.; Tashiro, K.; Nakamura, T.; Shiokawa, K.
Mech. Dev. 49, 123-131, 1995
A>Title: Molecular cloning of Xenopus HGF cDNA and its expression studies in Xenopus ea
A;Reference number: I51283; MUUD:95267690; PMID:7748783
A;Accession: I51283
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A;Residues: 1-710 <NA>
A;Cross-references: GB:S77422; NID:G989932; PIDN:AAB34354.1; PID:G989933
A>Note: the authors' translation for residue 458 (Thr) is inconsistent with the nucleot
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Function:
A>Description: stimulates mitosis of hepatocytes and other cells
A>Note: does not have protease activity
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: duplication; glycoprotein; growth factor; heterodimer; kringle
F;42-477, 478-709/Product: hepatocyte growth factor #status predicted <MAT>
F;42-477/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F;115-193/Domain: kringle homology <KR1>
F;198-275/Domain: kringle homology <KR2>
F;289-367/Domain: kringle homology <KR3>
F;375-453/Domain: kringle homology <KR4>
F;478-709/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
F;478-700/Domain: trypsin homology <TRY>
F;52,128,281,322,379,550,637,666/Binding site: carbohydrate (Asn) (covalent) #status pr
F;470-598/disulfide bonds: #status predicted

Query Match 23.0%; Score 127.5; DB 1; Length 710;

```

Best Local Similarity 30.1%; Pred. No. 9.1e-06;
Matches 31; Conservative 10; Mismatches 35; Indels 27; Gaps 5;

1 KTCYEGNGHFRGKASTDTWGRPCLPWNSATVLQQTTHAHR-----SDALQLGLGKH 52
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 287 KDCMKGGEGYRGVSVTYNGIQCRWDS-----QPPHLNFTPNYKCKDL-----SE 335

53 NYCNPNDNRPRWCY-----VQVGLKPLVOECMV-----HDCADG 87
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 336 NYCNPNGSESWCFTTDPNIRIRGHSQIKKQASNQOEYCG 378
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: May 25, 2004, 14:58:37
Job time : 6.55302 secs

GenCore version 5.1.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:43:40 ; Search time 3.70202 Seconds

(without alignment)
1350.274 Million cell updates/sec

Title: US-09-880-503-9

Perfect score: 554

Sequence: 1 KTCYCGNGHFGYRGKASTDTM.....QECMVHDCADGKKPSPEE 96

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	554	100.0	431	1 UROK_HUMAN	P00749 homo sapien
2	519	93.7	433	1 UROK_PAPCY	P16227 papio cynoc
3	437.5	79.0	442	1 UROK_PIG	P04188 sus scrofa
4	427	77.1	432	1 UROK_RAT	P29598 rattus norv
5	422	76.2	433	1 UROK_BOVIN	P05589 bos taurus
6	408	73.6	433	1 UROK_MOUSE	P06869 mus musculu
7	241	43.5	431	1 URT2_DESRO	P98121 desmodus ro
8	241	43.5	477	1 URT2_DESRO	P15638 desmodus ro
9	228.5	41.2	559	1 TPA_RAT	P19637 rattus norv
10	226	40.8	562	1 TPA_HUMAN	P00750 homo sapien
11	220	39.7	394	1 URT2_DESRO	P49150 desmodus ro
12	219	39.5	559	1 TPA_MOUSE	P11214 mus musculu
13	213	38.4	477	1 URT1_DESRO	P98119 desmodus ro
14	209	37.7	566	1 TPA_BOVIN	Q28198 bos taurus
15	198.5	35.8	434	1 UROK_CHICK	P15120 gallus gall
16	194.5	35.1	655	1 HGFA_HUMAN	Q04756 homo sapien
17	191.5	34.6	603	1 FA12_CAVPO	Q04962 cavia porce
18	188	33.9	653	1 HGPA_MOUSE	Q90988 mus musculu
19	170.5	30.8	615	1 FA12_HUMAN	P00748 homo sapien
20	161	29.1	1420	1 APOA_YACMU	P14417 macaca mula
21	159	28.7	473	1 KRMI_MOUSE	Q92484 rattus norv
22	159	28.7	473	1 KRMI_XENLA	Q90Y90 xenopus lae
23	158	28.5	452	1 KRMI_HUMAN	P08519 homo sapien
24	157	28.3	4548	1 APOA_HUMAN	Q96mu8 homo sapien
25	156	28.2	475	1 KRMI_HUMAN	P98140 bos taurus
26	155.5	28.1	593	1 FA12_BOVIN	P06868 bos taurus
27	152	27.4	812	1 PLMN_BOVIN	Q01177 rattus norv
28	150	27.1	169	1 PLMN_RAT	P12545 macaca mula
29	149	26.9	810	1 PLMN_YACMU	Q8ncw0 homo sapien
30	148.5	26.8	462	1 KRMI_HUMAN	P29627 homo sapien
31	148.5	26.8	711	1 HGFL_HUMAN	Q29485 erinaceus e
32	148.5	26.8	810	1 PLMN_ERIEU	P80009 canis famil
33	148	26.7	333	1 PLMN_CANFA	

34	148	26.7	790	1	PLMN_PIG	P06867 sus scrofa
35	147.5	26.6	716	1	HGFL_MOUSE	P26928 mus musculu
36	147.5	26.6	943	1	ROR2_HUMAN	Q01974 homo sapien
37	147.5	26.6	944	1	KRM2_MOUSE	Q92138 mus musculu
38	146.5	26.4	461	1	KRM2_MOUSE	Q8k187 mus musculu
39	146.5	26.4	810	1	PLMN_HUMAN	P00747 homo sapien
40	146	26.4	937	1	ROR1_HUMAN	Q01973 homo sapien
41	146	26.4	937	1	ROR1_MOUSE	Q92139 mus musculu
42	144.5	26.1	812	1	PLMN_MOUSE	P20918 mus musculu
43	139	25.1	685	1	ROR1_DROME	Q24488 drosophila
44	137.5	24.8	728	1	HGF_MOUSE	Q08048 mus musculu
45	135.5	24.5	728	1	HGF_RAT	P17945 rattus norv
46	133.5	24.1	728	1	HGF_HUMAN	P14210 homo sapien
47	131	23.6	325	1	PLMN_PETMA	P13574 petromyzon
48	125.5	22.7	618	1	THRB_MOUSE	P19221 mus musculu
49	124.5	22.5	343	1	PLMN_SHEEP	P81286 ovis aries
50	124	22.4	622	1	THRB_HUMAN	P00734 homo sapien
51	119.5	21.6	724	1	ROR2_DROME	Q9v6k3 drosophila
52	113	20.4	617	1	THRB_RAT	P18232 rattus norv
53	112	20.2	625	1	THRB_BOVIN	P00735 bos taurus
54	111	20.0	338	1	PLMN_HORSE	P80010 equus cabal
55	96	17.3	875	1	NETR_HUMAN	P56730 homo sapien
56	89	16.1	761	1	NETR_MOUSE	Q08782 mus musculu
57	73	13.2	559	1	ECM1_MOUSE	Q61508 mus musculu
58	70.5	12.7	540	1	ECM1_HUMAN	Q16610 homo sapien
59	69	12.5	369	1	V181_FOPPV	Q91552 fowlbox vir
60	68	12.3	161	1	ASF1_HELAN	P22357 helianthus
61	66	11.9	2871	1	FBN1_PIG	Q9tv36 sus scrofa
62	65.5	11.8	1278	1	NPC1_MOUSE	Q35604 mus musculu
63	65.5	11.8	1426	1	RHSD_ECOLI	P16919 escherichia
64	65	11.7	542	1	LNT_CHLMU	Q6pjx8 chlamydia m
65	63.5	11.5	542	1	LNT_CHLPR	O84539 chlamydia t
66	63.5	11.5	775	1	AD28_HUMAN	Q9ukq2 homo sapien
67	63	11.4	410	1	HMH2_DROME	P10035 drosophila
68	63	11.4	1342	1	Z335_HUMAN	Q9h422 homo sapien
69	62	11.2	1088	1	PRH_ETCTR	P48786 petroselinu
70	62	11.2	2871	1	FBN1_BOVIN	P98133 bos taurus
71	61	11.0	934	1	CLTC_HUMAN	P11586 h c-i-tetra
72	60	10.8	230	1	RT07_MARPO	P26867 marchantia
73	59.5	10.7	138	1	PA2A_VIPAA	P00626 vipera ammo
74	59.5	10.7	138	1	PA2B_VIPAA	P14424 vipera ammo
75	59.5	10.7	138	1	PA2C_VIPAA	P11407 vipera ammo

ALIGNMENTS

RESULT 1
UROK_HUMAN STANDARD; PRT; 431 AA.
ID UROK_HUMAN Q15844; Q16618; Q969W6;
AC P00749; O15844; Q16618; Q969W6;
DT 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLA0.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A.
ME MEDLINE=85215647; PubMed=2987867;
RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Biasi F.;
RT "The human, urokinase-plasminogen activator gene and its promoter.";
RL Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
RS SEQUENCE FROM N.A.
ME MEDLINE=85215647; PubMed=2987867;
RA Holmes W.B., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RT Steffens G.J., Heyneker H.L.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia coli.";

RA Biotechnology 3:923-929(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86056954; PubMed=2415429;
 RA Nagai M., Hiratsutsu R., Kaneda T., Hayasuke N., Arimura H.,
 RA Nishida M., Suyama T.;
 RT "Molecular cloning of cDNA coding for human preprourokinase.";
 RL Gene 36:183-188(1985).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85203359; PubMed=3888571;
 RA Jacobs P., Cravador A., Lorian R., Brockly F., Colau B., Chuchana P.,
 RA van Elsen A., Harzog A., Bollen A.;
 RT "Molecular cloning, sequencing, and expression in *Escherichia coli* of
 RT human preprourokinase cDNA.";
 RL DNA 4:139-146(1985).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85203359; PubMed=3888571;
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RL [6]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Lung;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Horg L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP SEQUENCE OF 66-431 FROM N.A.
 RX MEDLINE=84272706; PubMed=6589620;
 RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.;
 RT "Identification and primary sequence of an unspliced human urokinase
 RT poly(A)+ RNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
 RN [8]
 RP SEQUENCE OF 21-177.
 RX MEDLINE=83055084; PubMed=6754569;
 RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
 RA Flohe L.;
 RT "The primary structure of high molecular mass urokinase from human
 RT urine. The complete amino acid sequence of the A chain.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
 RN [9]
 RP SEQUENCE OF 156-176 AND 179-224.
 RX MEDLINE=83003608; PubMed=6749491;
 RA Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W.,
 RA Studer R.O.;
 RT "Human low-molecular-weight urinary urokinase. Partial
 RT characterization and preliminary sequence data of the two polypeptide
 RT chains.";
 RL Eur. J. Biochem. 125:251-257(1982).
 RN [10]
 RP SEQUENCE OF 158-410.
 RX MEDLINE=83055099; PubMed=6754572;
 RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;
 RT "The complete amino acid sequence of low molecular mass urokinase
 RT from human urine.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS)
 RX MEDLINE=96000858; PubMed=8591045;
 RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
 RA Dobson C.M., Stuart D.I., Jones E.Y.;
 RT "The crystal structure of the catalytic domain of human
 RT urokinase-type plasminogen activator.";
 RL Structure 3:681-691(1995).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
 RX MEDLINE=20266327; PubMed=10805774;
 RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
 RA Bode W., Magdolen V., Huber R., Moroder L.;
 RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
 RT selective inhibitors of human urokinase";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
 RN [13]
 RP STRUCTURE BY NMR.
 RX MEDLINE=89127526; PubMed=2536903;
 RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;
 RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
 RT dimensional NMR.";
 RL Nature 337:579-582(1989).
 RN [14]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=93003110; PubMed=1327118;
 RA Li X., Smith R.A.G., Dobson C.M.;
 RT "Sequential 1H NMR assignments and secondary structure of the kringle
 RT domain from urokinase.";
 RL Biochemistry 31:9562-9571(1992).
 RN [15]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=94149701; PubMed=8107091;
 RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
 RT "Solution structure of the kringle domain from urokinase-type
 RT plasminogen activator.";
 RL J. Mol. Biol. 235:1548-1559(1994).
 RN [16]
 RP VARIANT LEU-141.
 RX MEDLINE=96186279; PubMed=8652631;
 RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
 RA Sawasaki Y., Hanada K.;
 RT "Characterization of single chain urokinase-type plasminogen
 RT activator with a novel amino-acid substitution in the kringle
 RT structure";
 RL Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17]
 RP VARIANT LEU-141.
 RX MEDLINE=97218551; PubMed=9065988;
 RA Conne B., Berczy M., Belin D.;
 RT "Detection of polymorphisms in the human urokinase-type plasminogen
 RT activator gene.";
 RL Thromb. Haemost. 77:434-435(1997).
 RN [18]
 RP ERRATUM.
 RX Conne B., Berczy M., Belin D.;
 RL Thromb. Haemost. 78:973-973(1997).
 RN [19]
 RP VARIANT LEU-141.
 RX MEDLINE=97337920; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmalheidt B., Trommler P., Hell W.,
 RA Kreutzburg S., Graeff H., Magdolen V.;
 RT "Mutational analysis of the genes encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer";
 RL Electrophoresis 18:686-689(1997).
 CC -!- FUNCTION: Potent plasminogen activator and is clinically used for
 CC therapy of thrombolytic disorders.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- SUBUNIT: Found in high and low molecular mass forms. Each consists

CC of two chains, A and B. The high molecular mass form contains a
CC long chain A. Cleavage occurs after residue 155 in the low
CC molecular mass form to yield a short A1 chain.
CC -!- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
CC in Pulmonary Embolism (PE) to initiate fibrinolysis.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X02419; CAA26268.1; -
CC EMBL; M15476; AAA61253.1; -
CC EMBL; D00244; BAA00175.1; -
CC EMBL; D11143; BAA01919.1; -
CC EMBL; X02760; CAA26535.1; -
CC EMBL; AF377330; AAK53822.1; -
CC EMBL; BC013575; AAH13575.1; -
CC EMBL; K03226; AAC97138.1; -
CC EMBL; K02286; AAA61252.1; -
CC EMBL; A21571; CAA01559.1; -
CC EMBL; A18397; CAA01390.1; -
CC PIR; A00931; UKHU.
CC PDB; 1KDU; 31-OCT-93.

Query Match 100.0%; Score 554; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. NO. 8.3e-55;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLTQYTHAHS DALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLTQYTHAHS DALQLGLGKHNYCRNPDN 127
Qy 61 RRRPWCYVQGLKPLVQSCVHDCADGKXPSSPPEE 96
Db 128 RRRPWCYVQGLKPLVQSCVHDCADGKXPSSPPEE 163

RESULT 2
UROK_PAPCY STANDARD; PRT; 433 AA.
AC P16227;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T.; Wang T.W.; Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
RT plasminogen activator.";
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Found in high and low molecular mass forms. Each consists
CC of two chains, A and B. The high molecular mass form contains a
CC long chain A. Cleavage occurs after residue 155 in the low
CC molecular mass form to yield a short A1 chain (By similarity).

CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X51935; CAA36200.1; -
CC PIR; S14687; UKBAY.
CC HSP; P00749; ILMW.
CC MEROPS; S01.231; -
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR008293; Pept_S1A_uPA.
CC InterPro; IPR001254; Peptidase_S1.
CC Pfam; PF00051; Kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PIRSF001144; Urk_plasm_act; 1.
CC PRINTS; PIRSF001144; CHYMOTRYPSIN.
CC PRINTS; P3C0018; KRINGLE.
CC ProDom; P3C00395; Kringle; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; FALSE_NEG.
CC PROSITE; PS50026; EGF_3; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00700; KRINGLE_2; 1.
CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activator; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 26 62 EGF-LIKE.
FT DOMAIN 69 150 KRINGLE.
FT DOMAIN 151 177 CONNECTING PEPTIDE.
FT DOMAIN 178 433 SERINE PROTEASE.
FT DISULFID 30 38 BY SIMILARITY.
FT DISULFID 32 50 BY SIMILARITY.
FT DISULFID 52 61 INTERCHAIN (BY SIMILARITY).
FT DISULFID 167 298 BY SIMILARITY.
FT DISULFID 208 224 BY SIMILARITY.
FT DISULFID 216 287 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match 93.7%; Score 519; DB 1; Length 433;
Best Local Similarity 94.8%; Pred. NO. 7.1e-51;
Matches 91; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLTQYTHAHS DALQLGLGKHNYCRNPDN 60
Db 67 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLTQYTHAHS DALQLGLGKHNYCRNPDN 126

```
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
| | | | | | | | | | | | | | | | | | | | | |
DB 127 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 162

RESULT 3
UROK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCB1_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
EL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RA Nagamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; X01648; CAA25806.1; -.
DR EMBL; X02724; CAA26511.1; -.
DR PIR; A00932; UKPG.
DR HSSP; P00749; 1KDU.
DR MEROPS; S01.231; -.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR008293; Pept_S1A_uPA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00053; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PIRSF; PIRSF001144; Urk_plasm_act; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
```

```
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 189 CONNECTING PEPTIDE.
FT DOMAIN 190 442 SERINE PROTEASE.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
FT DISULFID 220 236 BY SIMILARITY.
FT DISULFID 228 299 BY SIMILARITY.
FT DISULFID 324 393 BY SIMILARITY.
FT DISULFID 356 372 BY SIMILARITY.
FT DISULFID 383 411 BY SIMILARITY.
FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
FT CONFLICT 241 241 Q -> H (IN REF. 1; CAA25806).
FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA26511).
FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
SQ SEQUENCE 442 AA; 49116 MW; EE32FCEP501321EE CRC64;

Query Match 79.0%; Score 437.5; DB 1; Length 442;
Best Local Similarity 74.3%; Pred. No. 1.1e-41;
Matches 78; Conservative 8; Mismatches 10; Indels 9; Gaps 1;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNSTVLTQTYHAHRSDALQLGLGKHNYCRPN 60
| | | | | | | | | | | | | | | | | | | | | |
DB 70 QTCFEGNGHSYRGKANTNTGRCPLPWNSTVLTQTYHAHRSDALQLGLGKHNYCRPN 129
| | | | | | | | | | | | | | | | | | | | | |
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
| | | | | | | | | | | | | | | | | | | | | |
DB 130 QRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 174
| | | | | | | | | | | | | | | | | | | | | |

RESULT 4
UROK_PIG STANDARD; PRT; 432 AA.
AC P29598;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCB1_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer 344;
RX MEDLINE=92233409; PubMed=1568219;
RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
RA Kefford R.F.;
RT "Transcriptional and posttranscriptional activation of urokinase
RT plasminogen activator gene expression in metastatic tumor cells.";
RL Cancer Res. 52:2489-2496(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Rabbani S.A.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
CC of two chains, A and B. The high molecular mass form contains a
CC long chain A. Cleavage occurs after residue 156 in the low
```

molecular mass form to yield a short A1 chain (By similarity).
-!- SIMILARITY: Belongs to peptidase family S1.
-!- SIMILARITY: Contains 1 EGF-like domain.
-!- SIMILARITY: Contains 1 kringle domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: X63434, CA945028.1; -
EMBL: X65651, CA94601.1; -
PIR: S24604, S18932.
HSSP: P00749, 1KDU.
MEROPS: S01.231; -
InterPro: IPR009003; Cys Ser trypsin.
InterPro: IPR006209; EGF-like.
InterPro: IPR006210; IEGF.
InterPro: IPR000001; Kringle.
InterPro: IPR008293; Pept_S1A_UA.
InterPro: IPR001254; Peptidase_S1.
InterPro: IPR001314; Peptidase_S1A.
Pfam: PF00051; kringle; 1.
Pfam: PF00089; trypsin; 1.
PIRSF: PIRSF001144; Urk plasem act; 1.
PRINTS: PR00722; CHYMOTRYPSIN.
PRINTS: PR00018; KRINGLE.
ProDom: PD000395; Kringle; 1.
SMART: SM00181; EGF; 1.
SMART: SM00130; KR; 1.
SMART: SM00020; Tryp_SPC; 1.
PROSITE: PS00022; EGF_1; 1.
PROSITE: PS01186; EGF_2; FALSE_NEG.
PROSITE: PS50026; EGF_3; 1.
PROSITE: PS00021; KRINGLE_1; 1.
PROSITE: PS00070; KRINGLE_2; 1.
PROSITE: PS00134; TRYPSIN_DOM; 1.
PROSITE: PS00240; TRYPSIN_HIS; FALSE_NEG.
PROSITE: PS00135; TRYPSIN_SRP; 1.
Plasminogen activator; Hydrolase; Serine protease; Glycoprotein; Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 20 177 CHAIN A (BY SIMILARITY).
FT CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 179 432 CHAIN B (BY SIMILARITY).
FT DOMAIN 27 63 EGF-LIKE.
FT DOMAIN 70 151 KRINGLE.
FT DOMAIN 152 178 CONNECTING PEPTIDE.
FT DOMAIN 179 432 SERINE PROTEASE.
FT DISULFID 31 39 BY SIMILARITY.
FT DISULFID 33 51 BY SIMILARITY.
FT DISULFID 53 62 BY SIMILARITY.
FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
FT DISULFID 210 226 BY SIMILARITY.
FT DISULFID 218 289 BY SIMILARITY.
FT DISULFID 314 383 BY SIMILARITY.
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 373 401 BY SIMILARITY.
FT ACT_SITE 225 225 CHARGE RELAY SYSTEM.
FT ACT_SITE 276 276 CHARGE RELAY SYSTEM.
FT ACT_SITE 377 377 CHARGE RELAY SYSTEM.
FT CONFLICT 16 24 E -> H (IN REF. 2).
FT CONFLICT 24 26 E -> G (IN REF. 2).
FT CONFLICT 332 332 D -> N (IN REF. 2).
FT SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Query Match 77.1%; Score 427; DB 1; Length 432;
Best Local Similarity 77.1%; Pred. No. 1.6e-40;
Matches 74; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYFGKASTDTNGRPCLPWNSATVLQOTYHHRSDALQGLGKHYCENPDN 60
Db 68 KTCYHGNGSGYRGKANTDTKGPCLAWNSFVLTQOTYHHRSDALSLGKHYCRNPDN 127
QY 61 RRRPCWYGVGLKPLVQECMVDGADGKPKSPPEE 96
Db 128 QRRPCWYGVGLKPLVQECMVDGADGKPKSPSTVDQ 163

RESULT 5

UROK_BOVIN STANDARD; PRT; 433 AA.
AC Q05589; Q28209;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aortic endothelium;
RX MEDLINE=93216119; PubMed=8385052;
RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,
RA Schleuning W.-D.;
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid.";
RL Gene 125:177-183(1993).
[2]
RP SEQUENCE OF 12-433 FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RL Int. Dairy J. 5:605-617(1995).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- INDUCTION: By retinoic acid.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L03546; AAA51419.1; -
CC EMBL: X85801; CA59796.1; -
CC PIR: JN0560; JN0560.
CC HSSP: P00749; 1LMW.
CC MEROPS: S01.231; -
CC InterPro: IPR009003; Cys Ser trypsin.
CC InterPro: IPR006209; EGF-like.
CC InterPro: IPR000001; Kringle.
CC InterPro: IPR008293; Pept_S1A_UA.
CC InterPro: IPR001254; Peptidase_S1.
CC InterPro: IPR001314; Peptidase_S1A.
CC Pfam: PF00051; kringle; 1.
CC Pfam: PF00089; trypsin; 1.
CC PIRSF: PIRSF001144; Urk plasmin act; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC PRINTS: PR00018; KRINGLE.
CC ProDom: PD000395; Kringle; 1.

DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 180 CONNECTING PEPTIDE.
FT DOMAIN 181 433 SERINE PROTEASE.
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 189 189 A -> T (IN REF. 2).
SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 76.2%; Score 422; DB 1; Length 433;
Best Local Similarity 75.0%; Pred. No. 5.7e-40;
Matches 72; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 KTCYEGNHPYRKASDTDMGRCLPWSATVLQQTVAHRSALQLGKGKNCNPDN 60
DB 70 KTCYQNGHSYRKARDLSGRCLAWDSFTVLLKWHARSDALQLGKGKNCNPDN 129

QY 61 RREPWCYVQGLKPLVQECMVHDCADKPKSPPEE 96
DB 130 QRREPCVQVGLKQFQVQFCMVQCSVKGSFSPREK 165

RESULT 6
UROK_MOUSE STANDARD; PRT; 433 AA.
AC P06869;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
GN PLAU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85179474; PubMed=2985383;
RA Belin D., Vassalli J.-D., Combepine C., Godeau F., Nagamine Y.,
RA Reich E., Koehler H.P., Duvoisin R.M.,
RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
RT mouse urokinase-type plasminogen activator.";
RL Eur. J. Biochem. 148:225-232 (1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88163489; PubMed=2831940;

RA RT Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;
RT "The murine urokinase-type plasminogen activator gene";
RL Biochemistry 26:8270-8279 (1987).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
CC of two chains, A and B. The high molecular mass form contains a
CC long chain A. Cleavage occurs after residue 156 in the low
CC molecular mass form to yield a short A1 chain (By similarity).
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X02389; CAA26231.1; -;
CC EMBL; M17922; AAA40539.1; -;
CC PIR; A29420; UKWS.
CC HSSP; P00749; 1KDU.
CC MEROPS; S01.231; -;
CC MGD; MGI:97611; Plau.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR008293; Pept_S1A_upa.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PIRSF001144; Urokinase; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; Kringle.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 178 CHAIN A (BY SIMILARITY).
FT CHAIN 157 178 SHORT A CHAIN (A1).
FT CHAIN 180 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 28 64 EGF-LIKE.
FT DOMAIN 71 152 KRINGLE.
FT DOMAIN 153 179 CONNECTING PEPTIDE.
FT DOMAIN 180 433 SERINE PROTEASE.
FT DISULFID 32 40 BY SIMILARITY.
FT DISULFID 34 52 BY SIMILARITY.
FT DISULFID 54 63 BY SIMILARITY.
FT DISULFID 169 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM.
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM.

FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
SQ SEQUENCE 433 AA; 48268 MW; A99C35F65043F9 CRC64;
Query Match 73.6%; Score 408; DB 1; Length 433;
Best Local Similarity 70.8%; Pred. No. 2.1e-38;
Matches 68; Conservative 12; Mismatches 16; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTPGRCPLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60
DB 69 KTCYEGNGSYRGKANTDTKGRPCLAWAPAVLQFYNAPDAISLGLGKHNYCRNPDN 128
QY 61 RRPWCYQVGLKPLVQECMVHDCADGKXSPSPPE 96
DB 129 QKRPWCYQIGLQRFVQECMVHDCSLSKKPSVVDQ 164
RESULT 7
UR2_DESRO STANDARD; PRT; 431 AA.
AC P98121;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA beta).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P., Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T., Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W., Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: Probably essential to support the feeding habits of this exclusively haematophagous animal. Probable potent thrombolytic agent.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M63989; AAA1594.1; -;
DR PIR; J50599; J50599.
DR HSSP; P98119; 1A51.
DR MEROPS; S01.239; -;
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.

DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00051; kringle_1.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD00018; KRINGLE.
DR ProDom; PD000395; Kringle_1.
DR SMART; SM00181; EGF_1.
DR SMART; SM00130; KR_1.
DR SMART; SM00020; Tryp_Spc_1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36 POTENTIAL.
FT CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.
FT DOMAIN 37 75 EGF-LIKE.
FT DOMAIN 82 163 KRINGLE.
FT DOMAIN 179 431 SERINE PROTEASE.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 41 52 BY SIMILARITY.
FT DISULFID 46 63 BY SIMILARITY.
FT DISULFID 65 74 BY SIMILARITY.
FT DISULFID 82 163 BY SIMILARITY.
FT DISULFID 103 145 BY SIMILARITY.
FT DISULFID 134 158 BY SIMILARITY.
FT DISULFID 168 299 BY SIMILARITY.
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 288 BY SIMILARITY.
FT DISULFID 313 388 BY SIMILARITY.
FT DISULFID 345 361 BY SIMILARITY.
FT DISULFID 378 406 BY SIMILARITY.
FT CARBOHYD 139 139 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 431 AA; 48221 MW; 699B5E675B162CBF CRC64;
Query Match 43.5%; Score 241; DB 1; Length 431;
Best Local Similarity 50.0%; Pred. No. 1.3e-19;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;
QY 2 TCYEGNGHFYRGKASTDTPGRCPLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPDN 61
DB 81 TCYKQGVYRGVTSSTSGAQCLNWNLSLTRTYNGRRSDAITLGLGHNHYCRNPDN 140
QY 62 RRPWCYQVGLKPLVQECMVHDC 85
DB 141 SKPWCYVTKASKFILEFCSPVCS 164
RESULT 8
UR2_DESRO STANDARD; PRT; 477 AA.
ID UR2_DESRO
AC P15638;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA alpha-2) (BAT-PA) (T-plasminogen activator).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;

[1] SEQUENCE FROM N.A.
 RP TISSUE=Salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat *Desmodus rotundus*: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN [2]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RX TISSUE=Salivary gland;
 RC MEDLINE=90036867; PubMed=2509450;
 RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
 RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
 RT "Isolation, characterization, and cDNA cloning of a vampire bat
 RT salivary plasminogen activator.";
 RL J Biol. Chem. 264:17947-17952(1989).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=9339059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of *Desmodus rotundus* (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 657:395-403(1992).
 CC -!- FUNCTION: Probably essential to support the feeding habits of this
 CC exclusively haematophagous animal. Probable potent thrombolytic
 CC agent.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- ENZYME REGULATION: Activity toward plasminogen is stimulated in
 CC the presence of fibrin I.
 CC -!- SUBUNIT: Monomer.
 CC -!- DOMAIN: The fibronectin type-I domain mediates binding to fibrin,
 CC and the kringle domain apparently mediates fibrin-induced
 CC stimulation of activity.
 CC -!- SIMILARITY: Belongs to peptidase family S1.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 1 kringle domain.

DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS01253; FIBONECTIN_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Signal; Multigene family.
 FT SIGNAL 1 36
 FT CHAIN 37 477
 FT DOMAIN 40 82
 FT DOMAIN 83 121
 FT DOMAIN 128 209
 FT DOMAIN 225 477
 FT ACT_SITE 272 272
 FT ACT_SITE 321 321
 FT ACT_SITE 428 428
 FT DISULFID 42 72
 FT DISULFID 70 79
 FT DISULFID 87 98
 FT DISULFID 92 109
 FT DISULFID 111 120
 FT DISULFID 128 209
 FT DISULFID 149 191
 FT DISULFID 180 204
 FT DISULFID 214 345
 FT DISULFID 257 273
 FT DISULFID 265 334
 FT DISULFID 359 434
 FT DISULFID 391 407
 FT DISULFID 424 452
 FT CARBOHYD 185 185
 FT CARBOHYD 398 398
 FT CONFLICT 403 403
 FT CONFLICT 417 417
 FT CONFLICT 435 435
 SQ SEQUENCE 477 AA; 53719 MW; 17486555C0E5077C CRC64;
 Query Match 43.5%; Score 241; DB 1; Length 477;
 Best Local Similarity 50.0%; Pred. No. 1.4e-19;
 Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;
 QY 2 TCYEGNGHFGKASTDTMGKPLPNSATVLCQYTHAHRSDALQGLGKHNYCRNPDR 61
 DB 127 TCYKQGVYTRGTWSTSSGAQCINWNSNLLTRTYNGRSDAITLGLGHNHNYCRNPDRN 186
 QY 62 RRPWCYVQGLKPLVQECWHDCA 85
 DB 187 SRPWCYVTKASKRFILEFCSPVCS 210
 RESULT 9
 TPA_RAT STANDARD; PRT; 559 AA.
 ID_TPA_RAT
 AC P19637;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
 DE (t-PA) (t-plasminogen activator).
 GN PLAT.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89170114; PubMed=3148445;

NV T., Leonardson G., Hsueh A.J.W.:
 "Cloning and characterization of a cDNA for rat tissue-type
 plasminogen activator."
 DNA 7:671-677(1988).
 [2]
 SEQUENCE FROM N.A.
 MEDLINE=90130448; PubMed=2105315;
 Feng P., Ohlsson M., NV T.:
 "The structure of the TATA-less rat tissue-type plasminogen activator
 gene. Species-specific sequence divergences in the promoter predict
 differences in regulation of gene expression."
 J. Biol. Chem. 265:2023-2027(1990).
 -!- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen
 to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By
 controlling plasmin-mediated proteolysis, it plays an important
 role in tissue remodeling and degradation, in cell migration and
 many other physiopathological events.
 -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 plasminogen to form plasmin.
 -!- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide
 bond.
 -!- SUBCELLULAR LOCATION: Secreted; extracellular.
 -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
 PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
 ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
 -!- MISCELLANEOUS: Binds to the kringle structure of the fibrin A
 chain. Binding to fibrin enhances its catalytic activity.
 -!- SIMILARITY: Belongs to peptidase family S1.
 -!- SIMILARITY: Contains 1 EGF-like domain.
 -!- SIMILARITY: Contains 1 fibronectin type I domain.
 -!- SIMILARITY: Contains 2 kringle domains.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 or send an email to license@isb-sib.ch.

 EMBL; M23697; AAA41812.1; --
 EMBL; M31197; AAA42261.1; --
 EMBL; M31185; AAA42261.1; JOINED.
 EMBL; M31186; AAA42261.1; JOINED.
 EMBL; M31187; AAA42261.1; JOINED.
 EMBL; M31188; AAA42261.1; JOINED.
 EMBL; M31189; AAA42261.1; JOINED.
 EMBL; M31190; AAA42261.1; JOINED.
 EMBL; M31191; AAA42261.1; JOINED.
 EMBL; M31192; AAA42261.1; JOINED.
 EMBL; M31193; AAA42261.1; JOINED.
 EMBL; M31194; AAA42261.1; JOINED.
 EMBL; M31195; AAA42261.1; JOINED.
 EMBL; M31196; AAA42261.1; JOINED.
 EMBL; A19618; CAA01482.1; --
 PIR; A35029; A35029.
 HSP; P00750; 1RTF.
 MEROPS; S01_232; --
 InterPro; IPR003003; Cys Ser trypsin.
 InterPro; IPR008209; EGF-like.
 InterPro; IPR000083; Fibrinctnl.
 InterPro; IPR006210; IEGF.
 InterPro; IPR000001; Kringle.
 InterPro; IPR001254; Peptidase S1.
 InterPro; IPR001314; Peptidase_S1A.
 Pfam; PF00008; EGF; 1.
 Pfam; PF00039; fn1; 1.
 Pfam; PF00051; kringle; 2.
 Pfam; PF00089; trypsin; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 PRINTS; PR00018; KRINGLE.
 ProDom; PD000395; kringle; 2.
 SMART; SM00181; KGF; 1.

GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in *E. coli*.";
RL Nature 301:214-221(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=8262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
RA Haiung N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BPV vectors.";
RL DNA 6:461-472(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=66196143; PubMed=3009482;
RA Frierer Degen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene";
RL J. Biol. Chem. 261:6972-6985(1986).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,
RA Opdenakker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in *Escherichia coli*.";
RL Mol. Biol. Med. 3:279-292(1986).
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Umbilical vein;
RX MEDLINE=90192129; PubMed=2107528;
RA Siebert P.D., Fong K.;
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
RT human endothelial cells";
RL Nucleic Acids Res. 18:1086-1086(1990).
RN [8]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Buetow K.H., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshikiyuki S., Carninci F., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

Bozak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Pahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [9]
RP SEQUENCE OF 212-361 FROM N.A.
RX MEDLINE=83169656; PubMed=6572897;
RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
RA Josephson S.;
RT "Isolation of cDNA sequences coding for a part of human tissue
RT plasminogen activator";
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
RN [10]
RP SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=85289338; PubMed=3161893;
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
RA Schleuning W.-D.;
RT "Isolation and characterization of the human tissue-type plasminogen
RT activator structural gene including its 5' flanking region.";
RL J. Biol. Chem. 260:11223-11230(1985).
RN [11]
RP SEQUENCE OF 31-562 FROM N.A.
RX MEDLINE=91291340; PubMed=1368681;
RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
RT "Purification and characterization of tissue plasminogen activator
RT secreted by human embryonic lung diploid fibroblasts, INR-90 cells.";
RL Agric. Biol. Chem. 55:1225-1232(1991).
RN [12]
RP SEQUENCE OF 36-562.
RC TISSUE=Melanoma;
RX MEDLINE=85000468; PubMed=6433976;
RA Pohl G., Kaelinstroem M., Bergsdorf N., Wallen P., Joernvall H.;
RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
RT derived amino acid sequence, identify the active site serine residue,
RT establish glycosylation sites, and localize variant differences.";
RL Biochemistry 23:3701-3707(1984).
RN [13]
RP SEQUENCE OF 33-52 AND 311-330.
RC TISSUE=Melanoma;
RX MEDLINE=83209620; PubMed=6682760;
RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
RT "Purification and characterization of a melanoma cell plasminogen
RT activator";
RL Eur. J. Biochem. 132:681-686(1983).
RN [14]
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=90092112; PubMed=2513186;
RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
RT "Carbohydrate structure of recombinant human uterine tissue
RT plasminogen activator expressed in mouse epithelial cells.";
RL Eur. J. Biochem. 186:273-286(1989).
RN [15]
RP CARBOHYDRATE-LINKAGE SITE THR-96.
RX MEDLINE=91159408; PubMed=1900431;
RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
RT "Tissue plasminogen activator has an O-linked fucose attached to
RT threonine-61 in the epidermal growth factor domain.";
RL Biochemistry 30:2311-2314(1991).
RN [16]
RP DISULFIDE BONDS IN KRINGLE 2.
RX MEDLINE=91244765; PubMed=1645336;
RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;
RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
RT plasminogen activator produced in *Escherichia coli*.";
RL J. Biol. Chem. 266:10070-10072(1991).

[17]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
RA Lambda D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
Rode W.;
RT "The 2.3 A crystal structure of the catalytic domain of recombinant
two-chain human tissue-type plasminogen activator.";
RL J. Mol. Biol. 258:117-135(1996).
[18]
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
RX MEDLINE=97449126; PubMed=9305622;
RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
Rode W.;
RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
crystal structure of single-chain human tPA.";
RL EMBO J. 16:4797-4805(1999).
[19]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
RX MEDLINE=92118803; PubMed=1310033;
RA de Vos A., Ultsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
Westbrook M.L., Kossakof A.A.;
RT "Crystal structure of the kringle 2 domain of tissue plasminogen
activator at 2.4-A resolution.";
RL Biochemistry 31:270-279(1992).
[20]
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=90122799; PubMed=2558718;
RA Byeon I.-J.L., Kelley R.F., Llinas M.;
RT "1H NMR structural characterization of a recombinant kringle 2 domain
from human tissue-type plasminogen activator.";
RL Biochemistry 28:9350-9360(1989).
[21]
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=91200042; PubMed=1901789;
RA Byeon I.-J.L., Kelley R.F., Llinas M.;
RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
assignments and secondary structure.";
RL Eur. J. Biochem. 197:155-165(1991).
[22]
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=92106329; PubMed=1762144;
RA Byeon I.-J.L., Llinas M.;
RT "Solution structure of the tissue-type plasminogen activator kringle
2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
drug.";
RL J. Mol. Biol. 222:1035-1051(1991).
[23]
Query Match 40.8%; Score 226; DB 1; Length 562;
Best Local Similarity 47.7%; Pred. No. 8.1e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
QY 2 TCYEGNGHYRGAASDTWGRCLPWSATVLOQTHAHRSDALQGLGKHNYCRPNDR 61
DB 126 TCYEDQGISYRGFTWSTAESGAECTNWNSSALAKPKYSGRRPDAIRLGLGNHNYCRPNDR 185
QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
DB 186 SKPWCYVFKAGKYSSSEFCSPACSEG 211
RESULT 11
URTG_DESGO
ID URTG_DESGO STANDARD; PRT; 394 AA.
AC P49150;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSFA
gamma).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Chiroptera; Microchiroptera; Phyllostomidae;

CC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
[1]
RP SEQUENCE FROM N.A.
RX TISSUE=Salivary gland;
MEDLINE=92039036; PubMed=1937019;
RA Knaetzel-Schmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
[2]
RP CHARACTERIZATION.
MEDLINE=93391059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
Knaetzel-Schmar J., Haendler B., Langer G., Baldus B., Witt W.,
Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: Probably essential to support the feeding habits of this
exclusively haematophagous animal. Probable potent thrombolytic
agent.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; M63990; AAA31595.1; -.
PIR; JS0600; JS0600.
DR MEROPS; P98119; IAS1.
DR HSSP; P98119; IAS1.
DR MEROPS; S01.239; -.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle_
InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
PFam; PF00051; kringle; 1.
PFam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
ProDom; PD000395; Kringle; 1.
SMART; SM00130; KR; 1; SPC; 1.
SMART; SM00020; Tryp_SPC; 1.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS00070; KRINGLE_2; 1.
PROSITE; PS00240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
Kringle; Signal; Multigene family.
KW SIGNAL 1 36 POTENTIAL.
FT CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
FT DOMAIN 45 126 KRINGLE.
FT DOMAIN 142 394 SERINE PROTEASE.
FT ACT_SITE 189 189 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 45 126 BY SIMILARITY.
FT DISULFID 66 108 BY SIMILARITY.
FT DISULFID 97 121 BY SIMILARITY.
FT DISULFID 131 262 BY SIMILARITY.
FT DISULFID 174 190 BY SIMILARITY.
FT DISULFID 182 251 BY SIMILARITY.
FT DISULFID 276 351 BY SIMILARITY.

[illegible]

FT	DISULFID	149	191	BY SIMILARITY.	
FT	DISULFID	180	204	BY SIMILARITY.	
FT	DISULFID	214	245		
FT	DISULFID	257	273		
FT	DISULFID	265	334		
FT	DISULFID	359	434		
FT	DISULFID	391	407		
FT	DISULFID	424	452	N-LINKED (GLCNAC. . .).	
FT	CARBOHYD	153	153	/FTID=CAR_000027. . .	
FT				N-LINKED (GLCNAC. . .).	
FT	CARBOHYD	398	398	/FTID=CAR_000028. . .	
FT	TURN	214	215		
FT	STRAND	223	224		
FT	TURN	226	227		
FT	STRAND	230	231		
FT	HELIX	234	236		
FT	TURN	238	239		
FT	STRAND	240	245		
FT	STRAND	254	263		
FT	TURN	264	265		
FT	STRAND	266	269		
FT	HELIX	271	273		
FT	TURN	280	282		
FT	STRAND	284	287		
FT	TURN	297	298		
FT	STRAND	300	309		
FT	TURN	311	312		
FT	TURN	315	317		
FT	TURN	319	320		
FT	STRAND	323	328		
FT	STRAND	338	338		
FT	TURN	339	340		
FT	STRAND	341	341		
FT	STRAND	345	345		
FT	TURN	349	350		
FT	TURN	355	356		
FT	STRAND	358	363		
FT	STRAND	366	366		
FT	STRAND	374	374		
FT	STRAND	379	385		
FT	STRAND	388	390		
FT	TURN	393	398		
FT	TURN	403	404		
FT	STRAND	405	409		
FT	TURN	425	426		
FT	TURN	428	429		
FT	STRAND	431	436		
FT	TURN	437	438		
FT	STRAND	439	448		
FT	TURN	455	456		
FT	STRAND	459	463		
FT	HELIX	464	467		
FT	HELIX	468	474		
SQ	SEQUENCE	477 AA;	53616 MW; AA06FD1739C10E5E CRC64;		
Query Match 38.4%; Score 213; DB 1; Length 477;					
Best Local Similarity 46.4%; Pred. No. 2e-16;					
Matches 39; Conservative 10; Mismatches 35; Indels 0; Gaps 0;					
QY	2 TCYEGNGHYRGKATDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGHNYCRNPDR 61				
DB	127 TCYEGQGVYRGWTSTAESRVEICINWSSLLTRTYNGRMPDAFNLGLGNHNYCRNPNGA 186				
QY	62 RRPWCYQVGLKPLVQEQMHDCA 85				
DB	187 KPWCVYIKAGKFTSSECSVPVCS 210				
RESULT 14					
TPA_BOVIN					
ID	TPA_BOVIN				STANDARD; PRT; 566 AA.
AC	Q28198;				

01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
15-MAR-2004 (Rel. 43, Last annotation update)
Tissue-type plasminogen activator precursor (EC 3.4.21.68) (TPA)
(t-PA) (t-plasminogen activator).
PLAT.
Bos taurus (Bovine).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
[1]
SEQUENCE FROM N.A.
Tissue=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
and tPA";
RL Int. Dairy J. 5:605-617(1995).
CC -!- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen
to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By
controlling plasmin-mediated proteolysis, it plays an important
role in tissue remodeling and degradation, in cell migration and
many other physiological events.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -!- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide
bond.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular.
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -!- MISCELLANEOUS: Binds to the kringle structure of the fibrin A
chain. Binding to fibrin enhances its catalytic activity.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 2 kringle domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X85800; CAA59795.1; .
HSP; P00750; 1RTP.
MEROPS; S01.232; .
InterPro; IPR009003; Cys_Ser_trypsin.
InterPro; IPR006209; EGF_like.
InterPro; IPR000083; Fibrinctnl.
InterPro; IPR006210; IEGF.
InterPro; IPR000001; Kringle.
InterPro; IPR001254; Peptidase_S1.
InterPro; IPR001314; Peptidase_S1A.
Pfam; PF00008; EGF; 1.
Pfam; PF00039; fnl; 1.
Pfam; PF00051; kringle; 2.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; KRINGLE.
ProDom; PD000395; Kringle; 2.
SMART; SM00181; EGF; 1.
SMART; SM00058; FN1; 1.
SMART; SM00130; KR; 2.
SMART; SM00020; Tryp_SPC; 1.
PROSITE; PS00022; EGF_1; 1.
PROSITE; PS01186; EGF_2; 1.
PROSITE; PS50026; EGF_3; 1.
PROSITE; PS01253; FIBRONECTIN_1; 1.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS50070; KRINGLE_2; 2.

DR	PROSITE; PS50240; TRYPSIN_DOM; 1.	DR	J. Biol. Chem. 265:1339-1344(1990).
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.	CC	-1- CATALYTIC ACTIVITY: Specific cleavage of Arg- -Val bond in
DR	PROSITE; PS00135; TRYPSIN_SER; 1.	CC	plasminogen to form plasmin.
KW	Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;	CC	-1- SIMILARITY: Belongs to peptidase family S1.
KW	Plasma; Kringle; EGF-like domain; Repeat; Signal.	CC	-1- SIMILARITY: Contains 1 EGF-like domain.
FT	SIGNAL 1 21	CC	-1- SIMILARITY: Contains 1 kringle domain.
FT	PROPEP 22 33	CC	-----
FT	CHAIN 34 344	CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration
FT	CHAIN 34 344	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
FT	CHAIN 34 344	CC	the European Bioinformatics Institute. There are no restrictions on its
FT	CHAIN 34 344	CC	use by non-profit institutions as long as its content is in no way
FT	CHAIN 34 344	CC	modified and this statement is not removed. Usage by and for commercial
FT	CHAIN 34 344	CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
FT	CHAIN 34 344	CC	or send an email to license@sib-sib.ch).
FT	CHAIN 34 344	CC	-----
FT	CHAIN 34 344	CC	EMBL; J05187; AAA49131.1; -
FT	CHAIN 34 344	CC	EMBL; J05188; AAA49130.1; -
FT	CHAIN 34 344	CC	PIR; A35005; A35005.
FT	CHAIN 34 344	CC	HSP; P00763; IDPO.
FT	CHAIN 34 344	CC	MEROPS; S01.231; -
FT	CHAIN 34 344	CC	InterPro; IPR009003; Cys Ser trypsin.
FT	CHAIN 34 344	CC	InterPro; IPR006209; EGF-like.
FT	CHAIN 34 344	CC	InterPro; IPR006210; EGF.
FT	CHAIN 34 344	CC	InterPro; IPR000001; Kringle.
FT	CHAIN 34 344	CC	InterPro; IPR008293; Pept_S1A_uPA.
FT	CHAIN 34 344	CC	InterPro; IPR001254; Peptidase S1.
FT	CHAIN 34 344	CC	InterPro; IPR001314; Peptidase_S1A.
FT	CHAIN 34 344	CC	Pfam; PF00051; kringle; 1.
FT	CHAIN 34 344	CC	Pfam; PF00089; trypsin; 1.
FT	CHAIN 34 344	CC	PIRSP; PIRSP001144; Urk_plasm act; 1.
FT	CHAIN 34 344	CC	PRINTS; PR00722; CHYMOTRYPSIN.
FT	CHAIN 34 344	CC	PRINTS; PR00018; KRINGLE.
FT	CHAIN 34 344	CC	PRODOM; PD000395; Kringle; 1.
FT	CHAIN 34 344	CC	SMART; SM00181; EGF; 1.
FT	CHAIN 34 344	CC	SMART; SM00130; KR; 1.
FT	CHAIN 34 344	CC	SMART; SM00020; TRYD_SPC; 1.
FT	CHAIN 34 344	CC	PROSITE; PS00022; EGF_1; 1.
FT	CHAIN 34 344	CC	PROSITE; PS01186; EGF_2; 1.
FT	CHAIN 34 344	CC	PROSITE; PS00026; EGF_3; 1.
FT	CHAIN 34 344	CC	PROSITE; PS00021; KRINGLE_1; 1.
FT	CHAIN 34 344	CC	PROSITE; PS00070; KRINGLE_2; 1.
FT	CHAIN 34 344	CC	PROSITE; PS50240; TRYPSIN_DOM; 1.
FT	CHAIN 34 344	CC	PROSITE; PS00134; TRYPSIN_HIS; 1.
FT	CHAIN 34 344	CC	PROSITE; PS00135; TRYPSIN_SER; 1.
KW	Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;	KW	Plasminogen activation; Hydrolase; Signal; Zymogen.
KW	Kringle; EGF-like domain; Signal; Zymogen.	KW	Kringle; EGF-like domain; Signal; Zymogen.
FT	SIGNAL 1 20	FT	POTENTIAL.
FT	CHAIN 21 434	FT	UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT	CHAIN 21 171	FT	CHAIN A (BY SIMILARITY).
FT	CHAIN 173 434	FT	CHAIN B (BY SIMILARITY).
FT	DOMAIN 36 72	FT	EGF-LIKE.
FT	DOMAIN 79 158	FT	KRINGLE.
FT	DOMAIN 159 172	FT	CONNECTING PEPTIDE.
FT	DOMAIN 173 434	FT	SERINE PROTEASE.
FT	DISULFID 40 48	FT	BY SIMILARITY.
FT	DISULFID 42 60	FT	BY SIMILARITY.
FT	DISULFID 62 71	FT	BY SIMILARITY.
FT	DISULFID 162 296	FT	INTERCHAIN (BY SIMILARITY).
FT	DISULFID 202 218	FT	BY SIMILARITY.
FT	DISULFID 210 285	FT	BY SIMILARITY.
FT	DISULFID 310 379	FT	BY SIMILARITY.
FT	DISULFID 342 358	FT	BY SIMILARITY.
FT	DISULFID 369 397	FT	BY SIMILARITY.
FT	ACT_SITE 217 217	FT	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 272 272	FT	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 273 373	FT	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	CARBOHYD 228 228	FT	N-LINKED (GLCNAC...) (POTENTIAL).
SQ	SEQUENCE 434 AA; 49400 MW; BD81048DD66A55 CRC64;	SQ	SEQUENCE 434 AA; 49400 MW; BD81048DD66A55 CRC64;
Query Match 35.8%; Score 198.5; DB 1; Length 434;			
Best Local Similarity 54.4%; Pred. No. 7.6e-15;			
Matches 37; Conservative 7; Mismatches 19; Indels 5; Gaps 2;			

```
QY 3 CYGNGHYRGKASTDTMGRPCLPMSATVLO-QTVHAHRSALQLGLGKHYCRNPDR 61
Db 79 CYSGNGEDYRGMAEDP-----GLYWDHPSVIRWGDYHADLKHALQLGLGKHYCRNPGR 134
QY 62 RRPWCYVQ 69
Db 135 SRPWCYTK 142

RESULT 16
HGFA_HUMAN STANDARD; PRT; 655 AA.
AC Q04756; Q14726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC TISSUE=Liver, and Serum;
RX MEDLINE=93252878; PubMed=7693665;
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
RA Kitamura N.;
RT "Molecular cloning and sequence analysis of the cDNA for a human
RT serine protease responsible for activation of hepatocyte growth
RT factor. Structural similarity of the protease precursor to blood
RT coagulation factor XII.";
RL J. Biol. Chem. 268:10024-10028(1993).
RN [2]
SEQUENCE OF 40-655 FROM N.A.
RA Zhao S., Odell C.;
RL Submitted (PEB-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Activates hepatocyte growth factor (HGF) by
CC converting it from a single chain to a heterodimeric form.
CC -!- SUBUNIT: Dimer of a short chain and a long chain linked by a
CC disulfide bond.
CC -!- SUBCELLULAR LOCATION: Secreted as an inactive single-chain
CC precursor and is then activated to a heterodimeric form.
CC -!- TISSUE SPECIFICITY: Liver.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- CAUTION: It is uncertain whether Met-1 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL; D14012; BAA03113.1; -.
DR EMBL; Z69923; -. NOT_ANNOTATED_CDS.
DR PIR; A46688; A46688.
DR MEROPS; S01.228; -.
DR Genew; HGNC:4894; HGFA.
DR MIM; 604552; -.
DR GO; GO:0005576; C:extracellular; TAS.
DR GO; GO:0004252; F:serine-type endopeptidase activity; TAS.
DR GO; GO:0006558; P:proteolysis and peptidolysis; TAS.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
```

```
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00083; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 30
FT PROPEP 31 372
FT CHAIN 373 407
FT CHAIN 408 655
FT DOMAIN 108 148
FT DOMAIN 160 198
FT DOMAIN 200 240
FT DOMAIN 241 279
FT DOMAIN 286 367
FT DOMAIN 408 655
FT ACT_SITE 447 447
FT ACT_SITE 497 497
FT ACT_SITE 598 598
FT DISULFID 108 133
FT DISULFID 122 148
FT DISULFID 164 175
FT DISULFID 169 186
FT DISULFID 188 197
FT DISULFID 202 230
FT DISULFID 228 237
FT DISULFID 245 256
FT DISULFID 267 278
FT DISULFID 286 367
FT DISULFID 307 349
FT DISULFID 338 362
FT DISULFID 394 521
FT DISULFID 432 448
FT DISULFID 440 510
FT DISULFID 535 604
FT DISULFID 567 583
FT DISULFID 594 622
FT CARBOHYD 48 48
FT CARBOHYD 290 290
FT CARBOHYD 468 468
FT CARBOHYD 492 492
FT CARBOHYD 546 546
FT CARBOHYD 644 644
FT CONFLICT 644 644
FT CHAIN 408 655
FT FIBRONECTIN TYPE-II.
FT EGF-LIKE 1.
FT FIBRONECTIN TYPE-I.
FT EGF-LIKE 2.
FT KRINGLE.
FT SERINE PROTEASE.
FT CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CHARGE RELAY SYSTEM (BY SIMILARITY).
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT INTERCHAIN (BY SIMILARITY).
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT R -> Q (IN REF. 2).
```

2 TCYEGNGHPYRGKASTDTMGRECLPWNSATVLQQTYHAHRSD-ALQLGLGKHNYCNPDN 60

Pfam; PF00040;	fn2: 1.
Pfam; PF00051;	kringle; 1.
Pfam; PF00089;	trypsin; 1.
PRINTS; PR00722;	CHYMOTRYPSIN.
PRINTS; PR00113;	ENTPEPIL.
PRINTS; PR00118;	KRINGLE.
ProDom; PD00095;	FN_Type_II; 1.
ProDom; PD000395;	Kringle; 1.
SMART; SMO0181;	EGF; 2.
SMART; SMO0059;	EN2; 1.
SMART; SMO0130;	KR; 1.
SMART; SMO0020;	Tryp_SPC; 1.
PROSITE; PS00022;	EGF_1; 2.
PROSITE; PS01186;	EGF_2; 1.
PROSITE; PS50036;	EGF_3; 2.
PROSITE; PS01253;	FIBONECTIN_1; 1.
PROSITE; PS00023;	FIBONECTIN_2; 1.
PROSITE; PS00021;	KRINGLE_1; 1.
PROSITE; PS50070;	KRINGLE_2; 1.
PROSITE; PS50240;	TRYPSIN_DOM; 1.
PROSITE; PS00134;	TRYPSIN_HIS; 1.
PROSITE; PS00139;	TRYPSIN_SER; 1.
Hydrolase; G1ycoprotein;	Plasma; Serine protease; Kringle; Signal;
EGF-like domain; Repeat;	Zymogen.
SIGNAL	1 29 BY SIMILARITY.
PROPEP	30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).
CHAIN	370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT CHAIN
CHAIN	406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG CHAIN.
DOMAIN	105 145 FIBONECTIN TYPE-II.
DOMAIN	157 195 EGF-LIKE 1.
DOMAIN	197 237 FIBONECTIN TYPE-I.
DOMAIN	238 276 EGF-LIKE 2.
DOMAIN	283 364 KRINGLE.
DOMAIN	406 653 SERINE PROTEASE.
ACT_SITE	445 445 CHARGE RELAY SYSTEM (BY SIMILARITY).
ACT_SITE	495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
ACT_SITE	596 596 CHARGE RELAY SYSTEM (BY SIMILARITY).
DISULFID	105 130 BY SIMILARITY.
DISULFID	119 145 BY SIMILARITY.
DISULFID	161 172 BY SIMILARITY.
DISULFID	166 183 BY SIMILARITY.
DISULFID	185 194 BY SIMILARITY.
DISULFID	199 227 BY SIMILARITY.
DISULFID	225 234 BY SIMILARITY.
DISULFID	242 253 BY SIMILARITY.
DISULFID	247 264 BY SIMILARITY.
DISULFID	266 275 BY SIMILARITY.
DISULFID	283 364 BY SIMILARITY.
DISULFID	304 346 BY SIMILARITY.
DISULFID	325 359 BY SIMILARITY.
DISULFID	392 519 INTERCHAIN (BY SIMILARITY).
DISULFID	430 446 BY SIMILARITY.
DISULFID	438 508 BY SIMILARITY.
DISULFID	533 602 BY SIMILARITY.
DISULFID	565 581 BY SIMILARITY.
DISULFID	592 620 BY SIMILARITY.
CARBOHYD	39 39 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD	47 47 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD	63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD	287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD	466 466 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD	544 544 N-LINKED (GLCNAC. .) (POTENTIAL).
CONFLICT	164 164 N-LINKED (GLCNAC. .) (POTENTIAL).
SEQUENCE	653 AA; 70567 MW; 88B4B20255DFDC CRC64; G -> W (IN REF. 2).
Query Match	33.9%; Score 188; DB 1; Length 653;
Best Local Similarity	53.0%; Pred. No. 1.8e-13;
Matches	35; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

```

Db      283 CFLGNTGTVRGVASTRAASGLSCLAWNSDLLYQELHVDVAAAVLLGLGPHAYCRNPKDE 342
QY      63 RWCYV 68
Db      343 RWCYV 348

RESULT 19
FA12_HUMAN
ID      FA12_HUMAN      STANDARD;      PRT;      615 AA.
AC      P00748; P78339;
DT      21-OCT-1986 (Rel. 01, Created)
DT      01-JUL-1989 (Rel. 12, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE      (HAF).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=88007593; PubMed=2888762;
RA      Cool D.E., McGillivray R.T.A.;
RT      "Characterization of the human blood coagulation factor XII gene.
RT      Intron/exon gene organization and analysis of the 5'-flanking
RT      region.";
RL      J. Biol. Chem. 262:13662-13673 (1987).
RN      [2]
RP      SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605.
RA      Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
RA      Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RA      Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE OF 4-615 FROM N.A.
RX      MEDLINE=86176794; PubMed=3754331;
RA      Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
RA      Cortese R.;
RT      "cDNA sequence coding for human coagulation factor XII (Hageman).";
RL      Nucleic Acids Res. 14:3146-3146 (1986).
RN      [4]
RP      SEQUENCE OF 14-615 FROM N.A.
RX      MEDLINE=86033830; PubMed=3877053;
RA      Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA      McGillivray R.T.A.;
RT      "Characterization of human blood coagulation factor XII cDNA.
RT      Prediction of the primary structure of factor XII and the tertiary
RT      structure of beta-factor XIIa.";
RL      J. Biol. Chem. 260:13666-13676 (1985).
RN      [5]
RP      SEQUENCE OF 146-615 FROM N.A.
RX      MEDLINE=86216049; PubMed=3011063;
RA      Que B.G., Davie E.W.;
RT      "Characterization of a cDNA coding for human factor XII (Hageman
RT      factor).";
RL      Biochemistry 25:1525-1528 (1986).
RN      [6]
RP      SEQUENCE OF 20-379.
RX      MEDLINE=85182674; PubMed=3886654;
RA      McMullen B.A., Fujikawa K.;
RT      "Amino acid sequence of the heavy chain of human alpha-factor XIIa
RT      (activated Hageman factor).";
RL      J. Biol. Chem. 260:5328-5341 (1985).
RN      [7]
RP      SEQUENCE OF 354-362 AND 373-615.
RX      MEDLINE=83291041; PubMed=6604055;
RA      Fujikawa K., McMullen B.A.;
RT      "Amino acid sequence of human beta-factor XIIa.";
RL      J. Biol. Chem. 258:10924-10933 (1983).
RN      [8]
RP      SEQUENCE OF 561-615 FROM N.A.
RC      TISSUE=Blood;

```

```

RX      MEDLINE=96133302; PubMed=8528215;
RA      Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammie B., Engel W.;
RT      "The novel acceptor splice site mutation 1139c(G->A) in the factor
RT      XII gene causes a truncated transcript in cross-reacting material
RT      negative patients.";
RL      Hum. Mol. Genet. 4:1235-1237 (1995).
RN      [9]
RP      CARBOHYDRATE-LINKAGE SITE THR-109.
RX      MEDLINE=92184750; PubMed=1544894;
RA      Harris R.J., Ling V.T., Spellman M.W.;
RT      "O-linked fucose is present in the first epidermal growth factor
RT      domain of factor XII but not protein C.";
RL      J. Biol. Chem. 267:5102-5107 (1992).
RN      [10]
RP      VARIANT WASHINGTON D.C. SER-590.
RX      MEDLINE=90046785; PubMed=2510163;
RA      Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B.,
RA      Saito H.;
RT      "Coagulation factor XII (Hageman factor) Washington D.C.: inactive
RT      factor XIIa results from Cys-571->Ser substitution.";
RL      Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322 (1989).
RN      [11]
RP      VARIANT LOCARNO PRO-372.
RX      MEDLINE=94325559; PubMed=8049433;
RA      Hovinga J.K., Schaller J., Stricker H., Willemin W.A., Furlan M.,
RA      Laemmle B.;
RT      "Coagulation factor XII Locarno: the functional defect is caused by
RT      the amino acid substitution Arg-353->Pro leading to loss of a
RT      kallikrein cleavage site.";
RL      Blood 84:1173-1181 (1994).
RN      [12]
RP      VARIANT TENRI CYS-53.
RX      MEDLINE=99290785; PubMed=10361128;
RA      Kondo S., Tokunaga F., Kawano S., Ono Y., Kumagai S., Koide T.;
RT      "Factor XII Tenri, a novel cross-reacting material negative factor XII
RT      deficiency, occurs through a proteasome-mediated degradation.";
RL      Blood 93:4300-4308 (1999).
CC      -!- FUNCTION: Factor XII is a serum glycoprotein that participates in
CC      the initiation of blood coagulation, fibrinolysis, and the
CC      generation of bradykinin and angiotensin.
CC      -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
CC      VII to form factor VIIa and factor XI to form factor Xla.
CC      -!- PTM: O- AND N-GLYCOSYLATED.
CC      -!- DISEASE: Defects in F12 do not cause any clinical symptoms. The
CC      sole effect is that whole-blood clotting time is prolonged.
CC      -!- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a
CC      complex bound to an anionic surface. Prekallikrein is cleaved by
CC      factor XII to form kallikrein, which then cleaves factor XII first
CC      to alpha-factor XIIa and then to beta-factor XIIa. Alpha-factor
CC      XIIa activates factor XI to factor Xla.
CC      -!- SIMILARITY: Belongs to peptidase family S1.
CC      -!- SIMILARITY: Contains 2 EGF-like domains.
CC      -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC      -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC      -!- SIMILARITY: Contains 1 kringle domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; M31315; AAA70225.1; -
DR      EMBL; AF538691; AAM57932.1; -
DR      EMBL; M11723; AAA51986.1; -
DR      EMBL; M17466; AAB59490.1; -
DR      EMBL; M17464; AAB59490.1; JOINED.
DR      EMBL; M17465; AAB59490.1; JOINED.
DR      EMBL; M13147; AAB70224.1; -
DR      EMBL; U71274; AAB51203.1; -
DR      PIR; A29411; KFHU12.

```


KW Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW Kringle; Repeat; Atherosclerosis.
FT NON_TER 1
FT DOMAIN 49 127 KRINGLE 1.
FT DOMAIN 163 241 KRINGLE 2.
FT DOMAIN 277 355 KRINGLE 3.
FT DOMAIN 391 469 KRINGLE 4.
FT DOMAIN 505 583 KRINGLE 5.
FT DOMAIN 619 697 KRINGLE 6.
FT DOMAIN 725 803 KRINGLE 7.
FT DOMAIN 839 917 KRINGLE 8.
FT DOMAIN 953 1031 KRINGLE 9.
FT DOMAIN 1067 1145 KRINGLE 10.
FT DOMAIN 1191 1420 SERINE PROTEASE.
SQ SEQUENCE 1420 AA; 158367 MW; BE10249E03C5B0E CRC64;

Query Match 29.1%; Score 161; DB 1; Length 1420;
Best Local Similarity 40.9%; Pred. No. 4.4e-10;
Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;

QY 3 CYEGNGHGYRGKASTDTMGRPCLPWNSATVLO--QTYHAHRS DALQLGLGKHNYCRNPDN 60
DB 1068 CYHNGQSRGTFSTVIGRTQCSHSMTPPHQKTPENHPNDLTM-----NYCRNPD 1122

QY 61 RRRPWCYVQGLKPLVQE--CMWHDCA 86
DB 1123 DTGPWCFT---MDPSVRREYCNLTRCSD 1147

RESULT 21
KRM1_MOUSE STANDARD; PRT; 473 AA.
ID KRM1_MOUSE
AC Q99N43;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringle-containing protein marking the eye and the nose) (Dickkopf receptor).
DE KREME1 OR KREME1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1] SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.
RP TISSUE=Brain, and Kidney;
RX MEDLINE=21167372; PubMed=11267660;
RA Nakamura T., Aoki S., Kitajima K., Takahashi T., Matsumoto K., Nakamura T.;
RT "Molecular cloning and characterization of Kremen, a novel kringle-containing transmembrane protein.";
RL Biochim. Biophys. Acta 1518:63-72(2001).
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: In the adult, widely expressed with high levels in heart, lung, kidney, skeletal muscle and testis.
CC -!- DEVELOPMENTAL STAGE: In the embryo, expression is first detected on day 9 and increases up to day 18. Lower levels are found in adult. At 9.5 dpc, expression is localised to the apical ectodermal ridge (AER) of the developing fore- and hindlimb buds, the telencephalon and the first brachial arch. At 10.5 dpc, expression is also observed in the myotome and in sensory tissues such as the nasal pit and optic vesicle.
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 WSC domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC ENBL; AB059617; BAB40968.1; -
DR HSP; P00747; ICEA.
DR MGD; MGI:1933988; Kremen.
DR GO; GO:0016021; C:Integral to membrane; NAS.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR002899; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00531; kringle; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
KW Wnt signaling pathway; Signal; Transmembrane; Kringle.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 473 KREMEN PROTEIN 1.
FT DOMAIN 21 392 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 393 413 POTENTIAL.
FT DOMAIN 414 473 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 114 KRINGLE.
FT DOMAIN 120 210 WSC.
FT DOMAIN 214 321 CUB.
FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 473 AA; 51716 MW; 586827788BE3FDD1 CRC64;

Query Match 28.7%; Score 159; DB 1; Length 473;
Best Local Similarity 45.8%; Pred. No. 2.3e-10;
Matches 31; Conservative 7; Mismatches 26; Indels 4; Gaps 2;

QY 3 CYEGNGHGYRGKASTDTM--GRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60
DB 32 CFTANGADYRGTSQSWTALQGKPCLFWNE--TFQHPYNTLYKYPNGEGGLGHEHNYCRNPDG 89

QY 61 RRRPWCYV 68
DB 90 DVSPWCYV 97

RESULT 22
KRM1_RAT STANDARD; PRT; 473 AA.
ID KRM1_RAT
AC Q924S4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringle-containing protein marking the eye and the nose) (Dickkopf receptor).
DE KREME1 OR KREME1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1] SEQUENCE FROM N.A.
RA Nakamura T., Nakamura T.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringle domain.

CC -1- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB065090; BAB62003.1; -
CC GO: GO:0016021; C: integral to membrane; ISS.
CC InterPro; IPR000859; CUB.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR002889; WSC.
CC Pfam; PF00431; CUB; 1.
CC Pfam; PF00051; Kringle; 1.
CC Pfam; PF01822; WSC; 1.
CC Pfam; PF00018; KR; 1.
CC PROSITE; PS01180; CUB; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.
CC Wnt signaling pathway; Signal; Transmembrane; Kringle.
CC SIGNAL 1 19
CC CHAIN 20 473 KREMEN PROTEIN 1.
CC DOMAIN 21 392 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 393 413 POTENTIAL.
CC DOMAIN 414 473 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 31 114 KRINGLE.
CC DOMAIN 120 210 WSC.
CC DOMAIN 214 321 CUB.
CC CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 473 AA; 51869 MW; 9B510857DF856F08 CRC64;
Query Match 28.7%; Score 159; DB 1; Length 473;
Best Local Similarity 45.6%; Pred. No. 2.3e-10;
Matches 31; Conservative 7; Mismatches 26; Indels 4; Gaps 2;
QY 3 CYEGNGHYRGKASTDTM--GRPCLPWNSATVLOQTHAHSRDLQLGLGKHNYCRNPDN 60
DB 32 CFTANGADYRGTSWTALQGGKPCLFWNE--TFQHPYNTLKYPNGEGGLGHEHNYCRNPDG 89
QY 61 RRRPWCYV 68
DB 90 DVSPWCYV 97
RESULT 23
ID_KRM1_XENLA STANDARD; PRT; 452 AA.
AC Q90Y90;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakamura T.;

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block wnt/beta-catenin signaling (by similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- SIMILARITY: Contains 1 CUB domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB070851; BAB64294.1; -
CC InterPro; IPR000859; CUB.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR002889; WSC.
CC Pfam; PF00431; CUB; 1.
CC Pfam; PF00051; Kringle; 1.
CC Pfam; PF01822; WSC; 1.
CC PRINTS; PR00018; KRINGLE.
CC PRODOM; PD000395; Kringle; 1.
CC SMART; SM00042; CUB; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00321; WSC; 1.
CC PROSITE; PS01180; CUB; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.
CC Wnt signaling pathway; Glycoprotein; Kringle;
CC Transmembrane; Signal.
CC SIGNAL 1 22
CC CHAIN 23 452 KREMEN PROTEIN 1.
CC DOMAIN 23 369 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 370 390 POTENTIAL.
CC DOMAIN 391 452 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 29 112 KRINGLE.
CC DOMAIN 118 208 WSC.
CC DOMAIN 212 319 CUB.
CC CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 452 AA; 50188 MW; ED24BCD1AF4564E2 CRC64;
Query Match 28.5%; Score 158; DB 1; Length 452;
Best Local Similarity 44.1%; Pred. No. 2.8e-10;
Matches 30; Conservative 9; Mismatches 25; Indels 4; Gaps 2;
QY 3 CYEGNGHYRGKASTDTM--GRPCLPWNSATVLOQTHAHSRDLQLGLGKHNYCRNPDN 60
DB 30 CYTVNGADYRGTSWTALQGGKPCLFWNE--TFQHPYNTLKYPNGEGGLGHEHNYCRNPDG 87
QY 61 RRRPWCYV 68
DB 88 DVSPWCYV 95
RESULT 24
ID_APOA_HUMAN STANDARD; PRT; 4548 AA.
AC P08519;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).
GN LPA.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89039109; PubMed=3670400; Eaton D.L., Chen E.Y.,
RA McLean J.W., Tomlison J.E., Kuang W.-J., Scannu A.M., Law R.M.,
RA Fless G.M., Scannu A.M., Law R.M.;
RT "cDNA sequence of human apolipoprotein(a) is homologous to
RL plasminogen.";
RL Nature 330:132-137(1987).
RN [2]
RP SERINE PROTEASE ACTIVITY.
RX MEDLINE=90076123; PubMed=2531657;
RA Salonen E.-M., Jauhainen M., Zardi L., Vaheri A., Ehnholm C.;
RT "Lipoprotein(a) binds to fibronectin and has serine proteinase
RL activity capable of cleaving it.";
RL EMBO J. 8:4035-4040(1989).
RN [3]
RP REVIEW.
RX MEDLINE=90049223; PubMed=2530631;
RA Utermann G.;
RT "The mysteries of lipoprotein(a).";
RL Science 246:904-910(1989).
RN [4]
RP CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.
RX MEDLINE=21303595; PubMed=11294842;
RA Garner B., Merry A.H., Royce L., Harvey D.J., Rudd P.M., Thillet J.;
RT "Structural elucidation of the N- and O-glycans of human
RL apolipoprotein(a): role of o-glycans in conferring protease
RT resistance.";
RL J. Biol. Chem. 276:22200-22208(2001).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.
RX MEDLINE=96217891; PubMed=8642595;
RA Mikol V., Lograsso P.V., Boettcher B.R.;
RT "Crystal structures of apolipoprotein(a) kringle IV/7 free and
RL complexed with 6-aminohexanoic acid and with p-aminomethylbenzoic
RL acid: existence of novel and expected binding modes.";
RL J. Mol. Biol. 256:751-761(1996).
RN [6]
RP VARIANT ARG-4193.
RX MEDLINE=95002201; PubMed=7918682;
RA Scannu A.M., Lee J.C., Hinman J.;
RT "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37
RL associated with a lysine binding defect in Lp(a).";
RL Biochim. Biophys. Acta 1227:41-45(1994).
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autoproteolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin.
CC -!- PTM: N- and O-glycosylated. The N-glycans are complex biantennary
CC structures present in either a mono- or disialylated state. The
CC O-glycans are mostly (80%) represented by the monosialylated core
CC type I structure, NeuAcGalNAc2-GalNAc6SialNAc, with smaller
CC amounts of disialylated and non-sialylated O-glycans also
CC detected.
CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its
CC naturally occurring proteolytic fragments are correlated with
CC atherosclerosis. Homology with plasminogen kringle IV and V is
CC thought to underlie the atherogenicity of the protein, because the
CC fragments are competing with plasminogen for fibrin(ogen) binding.
CC -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC leading to the formation of the so called mini-Lp(a). Apo(a)
CC fragments accumulate in atherosclerotic lesions, where they may
CC promote thrombogenesis. O-glycosylation may limit the extent of
CC proteolytic fragmentation.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -!- SIMILARITY: Contains 38 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X06290; CAA29618.1; --
CC PIR; S00857; S00857.
CC PDB; 1171; 13-JUN-01.
CC PDB; 1JFN; 28-JUN-02.
CC PDB; 1KIV; 18-MAY-99.
CC PDB; 3KIV; 18-MAY-99.
CC PDB; 4KIV; 18-MAY-99.
CC MEROPS; S01.226; --
CC Genew; HGNC:6667; LPA.
CC MIM; 152200; --
CC GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
CC GO; GO:0008015; P:circulation; TAS.
CC GO; GO:0009405; P:pathogenesis; TAS.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000001; Kringle_Sl.
CC InterPro; IPR001254; Peptidase_Sl.
CC InterPro; IPR001314; Peptidase_SlA.
CC Pfam; PF00051; Kringle; 38.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC PRODOM; PD000395; Kringle; 38.
CC SMART; SM00130; KR; 38.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00021; KRINGLE_1; 38.
CC PROSITE; PS00070; KRINGLE_2; 38.
CC PROSITE; PS02040; TRYPsin_DOM; 1.
CC PROSITE; PS00134; TRYPsin_HIS; 1.
CC PROSITE; PS00135; TRYPsin_SER; 1.
CC HydroLase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW Kringle; Repeat; Atherosclerosis; Signal; Polymorphism; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 4548 APOLIPOPROTEIN(A)
FT DOMAIN 20 130 KRINGLE TYPE IV, 1.
FT DOMAIN 131 244 KRINGLE TYPE IV, 2.
FT DOMAIN 245 358 KRINGLE TYPE IV, 3.
FT DOMAIN 359 472 KRINGLE TYPE IV, 4.
FT DOMAIN 473 586 KRINGLE TYPE IV, 5.
FT DOMAIN 587 700 KRINGLE TYPE IV, 6.
FT DOMAIN 701 814 KRINGLE TYPE IV, 7.
FT DOMAIN 815 928 KRINGLE TYPE IV, 8.
FT DOMAIN 929 1042 KRINGLE TYPE IV, 9.
FT DOMAIN 1043 1156 KRINGLE TYPE IV, 10.
FT DOMAIN 1157 1270 KRINGLE TYPE IV, 11.
FT DOMAIN 1271 1384 KRINGLE TYPE IV, 12.
FT DOMAIN 1385 1498 KRINGLE TYPE IV, 13.
FT DOMAIN 1499 1612 KRINGLE TYPE IV, 14.
FT DOMAIN 1613 1726 KRINGLE TYPE IV, 15.
FT DOMAIN 1727 1840 KRINGLE TYPE IV, 16.
FT DOMAIN 1841 1954 KRINGLE TYPE IV, 17.
FT DOMAIN 1955 2068 KRINGLE TYPE IV, 18.
FT DOMAIN 2069 2182 KRINGLE TYPE IV, 19.
FT DOMAIN 2183 2296 KRINGLE TYPE IV, 20.
FT DOMAIN 2297 2410 KRINGLE TYPE IV, 21.
FT DOMAIN 2411 2524 KRINGLE TYPE IV, 22.
FT DOMAIN 2525 2638 KRINGLE TYPE IV, 23.
FT DOMAIN 2639 2752 KRINGLE TYPE IV, 24.
FT DOMAIN 2753 2866 KRINGLE TYPE IV, 25.
FT DOMAIN 2867 2980 KRINGLE TYPE IV, 26.
FT DOMAIN 2981 3094 KRINGLE TYPE IV, 27.
FT DOMAIN 3095 3208 KRINGLE TYPE IV, 28.
FT DOMAIN 3209 3322 KRINGLE TYPE IV, 29.
FT DOMAIN 3323 3436 KRINGLE TYPE IV, 30.
FT DOMAIN 3437 3550 KRINGLE TYPE IV, 31.
FT DOMAIN 3551 3664 KRINGLE TYPE IV, 32.
FT DOMAIN 3665 3770 KRINGLE TYPE IV, 33.

FT DOMAIN 3771 3884 KRINGLE TYPE IV, 34.
FT DOMAIN 3885 3998 KRINGLE TYPE IV, 35.
FT DOMAIN 3999 4112 KRINGLE TYPE IV, 36.
FT DOMAIN 4113 4226 KRINGLE TYPE IV, 37.
FT DOMAIN 4227 4340 KRINGLE TYPE V.
FT DOMAIN 4341 4454 SERINE PROTEASE.
FT ACT_SITE 4328 4369 CHARGE RELAY SYSTEM.
FT ACT_SITE 4412 4452 CHARGE RELAY SYSTEM.
FT ACT_SITE 4498 4498 CHARGE RELAY SYSTEM.
FT VARIANT 4193 4193 W -> R (LOSS OF LYSINE-SEPHAROSE BINDING).
SQ SEQUENCE 4548 AA; 501313 MW; 969215E96A465CSF CRC64;
Query Match 28.3%; Score 157; DB 1; Length 4548;
Best Local Similarity 34.2%; Pred. No. 4.2e-09;
Matches 40; Conservative 9; Mismatches 36; Indels 32; Gaps 6;
Qy 3 CYEENGHYRGKASTDMGRCPLPWNATVLOQTYHAHRSAL---QLGLGKHNYCRNP 59
Db 3896 CYRGDGSYRGTLSTTIGRTGCSWS-----MTPWHRRIRPLYPNAGLTR-NYCRNP 3949
Qy 60 NRBRPCVQVGLKPLV--CEQVHVC-----ADGKKPSPPEE 96
Db 3950 ABIRPWCYT---MDPSVRWBYCNLTRCPVTSESVLTPTVAPVPSTEAPSEQAPPEK 4003
RESULT 25
KEMI_HUMAN
ID KEMI_HUMAN STANDARD; PRT; 475 AA.
AC Q96M08; Q96Y70; Q96G55; Q96U1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (kringle-containing protein marking the eye and the nose) (Dickkopf receptor).
GN KREMEN1 OR KREMEN
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Nakamura T., Nakamura T.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Taghito H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Maseuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Nagasuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K.,
RA Isogai T.;
RL "NEDO human cDNA sequencing project";
RN Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RP [3]
RX SEQUENCE FROM N.A. PubMed=10591208;
RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,

RA Martyn J.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillips B.C.T.,
RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawabaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuayama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Hu Y., Hu A., Kanton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qian Y., Ray L., Ren O., Shaull D., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoso S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozerky P., Rohlfing T.,
RA Scheet P., Walker C., Wamley A., Wohldmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelmann L., Kim U.-J., Shizuya H., Simon M.I., Dumanski J.P.,
RA Peyrard M., Kedra D., Seroussi E., Franconi I., Tapia I., Bruder C.E.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tilahun Y., Wright H.;
RT "The DNA sequence of human chromosome 22";
RL Nature 402:489-495 (1999).
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q96MU8-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q96MU8-2; Sequence=VSP_003900;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB059618; BAB40969.1; -;
CC EMBL; AK056425; BAB71180.1; -;
CC EMBL; Z95116; CAB62952.1; -;
CC EMBL; AL021393; CAB62959.1; -;
CC GenSeq; HGNC:17550; KREMEN1.
CC GO; GO:0016021; C:integral to membrane; ISS.
CC GO; GO:0005624; C:membrane fraction; TAS.
CC GO; GO:0007154; P:cell communication; TAS.
CC InterPro; IPR000859; CUB.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR002889; WSC.
CC Pfam; PF00431; CUB; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF01822; WSC; 1.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00042; CUB; 1.
CC SMART; SM00130; KR; 1.
CC PROSITE; PS01180; CUB; 1.
CC PROSITE; PS00021; KRINGLE; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.

FT	DISULFID	88	100	BY SIMILARITY.
FT	DISULFID	94	109	BY SIMILARITY.
FT	DISULFID	111	120	BY SIMILARITY.
FT	DISULFID	125	153	BY SIMILARITY.
FT	DISULFID	151	160	BY SIMILARITY.
FT	DISULFID	168	179	BY SIMILARITY.
FT	DISULFID	173	188	BY SIMILARITY.
FT	DISULFID	190	199	BY SIMILARITY.
FT	DISULFID	207	287	BY SIMILARITY.
FT	DISULFID	230	269	BY SIMILARITY.
FT	DISULFID	258	282	BY SIMILARITY.
FT	DISULFID	336	463	BY SIMILARITY.
FT	DISULFID	374	390	BY SIMILARITY.
FT	DISULFID	382	452	BY SIMILARITY.
FT	DISULFID	413	416	BY SIMILARITY.
FT	DISULFID	479	547	BY SIMILARITY.
FT	DISULFID	510	526	BY SIMILARITY.
FT	DISULFID	537	568	BY SIMILARITY.
FT	CARBOHYD	99	99	O-LINKED (FUC) (BY SIMILARITY).
FT	CARBOHYD	241	241	N-LINKED (GLNAC. .) (POTENTIAL).
FT	CARBOHYD	263	263	N-LINKED (GLNAC. .) (POTENTIAL).
FT	CARBOHYD	410	410	N-LINKED (GLNAC. .) (POTENTIAL).
SO	SEQUENCE	593 AA; 65148 MW; 721592BA792BD61F CRC64;		
Query Match 28.1%; Score 155.5; DB 1; Length 593;				
Best Local Similarity 38.4%; Pred. No. 7.2e-10;				
Matches 33; Conservative 9; Mismatches 37; Indels 7; Gaps 3;				
QY	2 TCYE--GNGHYRCKASTDWMRCPLEWNSATVLCQTY-HARRSDALQLGLGHVYCRNP 58			
Db	206 SCYDDRDRGLSYRGMAGTTTSGAPCQSWAS-----EATYMNVTAEQVLNGLGDFHAFCRNP 261			
QY	59 DNRRRPWCYQVGLKPLVQECVHDC 84			
Db	262 DNDTRPWCFFIWKGRDLSWNYCELAPC 287			
RESULT 27				
PLMN	BOVIN	STANDARD;	PRT;	812 AA.
AC	P06858; Q28162;			
DT	01-JAN-1988 (Rel. 06, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DE	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Plasminogen precursor (EC 3.4.21.7).			
GN	PLG.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RA	Berglund L., Andersen M.D., Petersen T.E.;			
RT	"Cloning and characterization of the bovine plasminogen cDNA.";			
RL	Int. Dairy J. 5:593-603(1995).			
RN	[2]			
RP	SEQUENCE OF 27-812, AND CARBOHYDRATE-LINKAGE SITES.			
RX	MEDLINE=85203906; PubMed=3846532;			
RA	Schaller J., Moser P.W., Dannegger-Muller G.A.K., Rosselet S.J.,			
RA	Kampfer U., Rickli E.E.;			
RT	"Complete amino acid sequence of bovine plasminogen. Comparison with			
RT	human plasminogen.";			
RL	Eur. J. Biochem. 149:267-278(1985).			
RN	[3]			
RP	SEQUENCE OF 706-812 FROM N.A.			
RX	MEDLINE=85023311; PubMed=6148961;			
RA	Malinowski D.P., Sadler J.E., Davie E.W.;			
RT	"Characterization of a complementary deoxyribonucleic acid coding for			
RT	human and bovine plasminogen.";			
RL	Biochemistry 23:4243-4250(1984).			
RN	[4]			
RP	CARBOHYDRATE-LINKAGE SITES.			
RX	MEDLINE=88185329; PubMed=3356193;			
RA	Marti T., Schaller J., Rickli E.E., Schmid K., Kamberling J.P.,			
RA	Gerwig G.J., van Halbeek H., Vliegthart J.F.;			
RT	"The N- and O-linked carbohydrate chains of human, bovine and porcine			
RT	plasminogen. Species specificity in relation to sialylation and			
RT	fucosylation patterns.";			
RL	Eur. J. Biochem. 173:57-63(1988)			
CC	-!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as			
CC	a proteolytic factor in a variety of other processes including			
CC	embryonic development, tissue remodeling, tumor invasion,			
CC	and inflammation; in ovulation it weakens the walls of the			
CC	Graafian follicle. It activates the urokinase-type plasminogen			
CC	activator, collagenases and several complement zymogens, such			
CC	as C1 and C5. It cleaves fibrin, fibronectin, thrombospondin,			
CC	laminin and von Willebrand factor.			
CC	-!- CATALYTIC ACTIVITY: Preferential cleavage: Lys- -Xaa > Arg- -Xaa;			
CC	higher selectivity than trypsin. Converts fibrin into soluble			
CC	products.			
CC	-!- ENZYME REGULATION: Converted into plasmin by plasminogen			
CC	activators, both plasminogen and its activator being bound to			
CC	fibrin. Cannot be activated with streptokinase.			
CC	-!- PTM: N-LINKED GLYCAN CONTAIN N-ACETYLACTOSAMINE AND SIALIC ACID.			
CC	O-LINKED GLYCANS CONSIST OF GAL-GALNAc DISACCHARIDE WITH IS			
CC	MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).			
CC	-!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin			
CC	immediately after dissociation from the clot.			
CC	-!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.			
CC	-!- SIMILARITY: Contains 5 kringle domains.			
CC	-----			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X79402; CAA55939.1; -			
DR	EMBL; K02935; AAA30714.1; -			
DR	PIR; S45046; PLBO.			
DR	HSSP; P00747; 2PK4.			
DR	MEROPS; S01.233; -			
DR	GlycoSuiteDB; P06868; -			
DR	InterPro; IPR009003; Cys_Ser_trypsin.			
DR	InterPro; IPR000001; Kringle.			
DR	InterPro; IPR003014; PAN.			
DR	InterPro; IPR003609; Pan_app.			
DR	InterPro; IPR001254; Peptidase S1.			
DR	InterPro; IPR001314; Peptidase_S1A.			
DR	InterPro; IPR003966; Peptidase_S1A_pr.			
DR	Pfam; PF00051; kringle; 5.			
DR	Pfam; PF00024; PAN; 1.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	PRINTS; PR00018; KRINGLE.			
DR	PRINTS; PR01505; PROTHROMBIN.			
DR	ProDom; PD000395; Kringle; 5.			
DR	SMART; SM00130; KR; 5.			
DR	SMART; SM00473; PAN_AP; 1.			
DR	SMART; SM00020; TRYP_SPC; 1.			
DR	PROSITE; PS00021; KRINGLE_1; 5.			
DR	PROSITE; PS00070; KRINGLE_2; 5.			
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;			
KW	tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;			
KW	Signal.			
FT	CHAIN	1	26	PLASMINOGEN.
FT	CHAIN	27	812	PLASMIN HEAVY CHAIN A.
FT	CHAIN	27	583	PLASMIN LIGHT CHAIN B.
FT	CHAIN	584	812	

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

inflammation; in ovulation it weakens the walls of the Graafian follicle. It activates the urokinase-type plasminogen activator, collagenase and several complement zymogens, such as C1 and C5. It cleaves fibrin, fibronectin, thrombospondin, laminin and von Willebrand factor.

-i- CATALYTIC ACTIVITY: Preferential cleavage: Lys-[Xaa > Arg]-Xaa; higher selectivity than trypsin. Converts fibrin into soluble products.

-i- ENZYME REGULATION: Converted into plasmin by plasminogen activators, both plasminogen and its activator being bound to fibrin. Activated with catalytic amounts of streptokinase.

-i- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin immediately after dissociation from the clot.

-i- MISCELLANEOUS: In the presence of the inhibitor, the activation involves only cleavage after Arg-580, resulting in 2 chains held together by 2 disulfide bonds. Without the inhibitor, the activation involves also removal of the activation peptide.

-i- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.

-i- SIMILARITY: Contains 5 kringle domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL: J04697; AAA36901.1; -
 PIR: B32869; B30848.
 HSRP: P00747; 1PWK.
 MEQSPS; S01.233; -
 InterPro: IPR003003; Cys_Ser_trypsin.
 InterPro: IPR000001; Kringle.
 InterPro: IPR003014; PAN.
 InterPro: IPR003609; Pan_app.
 InterPro: IPR001254; Peptidase_S1.
 InterPro: IPR001314; Peptidase_S1A.
 Pfam: PF00051; kringle; 5.
 Pfam: PF00024; PAN; 1.
 Pfam: PF00089; trypsin; 1.
 PRINTS: PR00722; CHYMOTRYPSIN.
 PRINTS: PR00018; KRINGLE.
 PRINTS: PR01505; PROTHROMBIN.
 ProDom: PD000395; kringle; 5.
 SMART: SM00130; KR; 4.
 SMART: SM00473; PAN AP; 1.
 PROSITE: PS00021; KRINGLE_1; 5.
 PROSITE: PS00070; KRINGLE_2; 5.
 PROSITE: PS0240; TRYPSIN_DOM; 1.
 PROSITE: PS00134; TRYPSIN_HIS; 1.
 PROSITE: PS00135; TRYPSIN_SER; 1.
 Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
 KW Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 810 PLASMINOGEN
 FT CHAIN 20 580 PLASMIN HEAVY CHAIN A.
 FT PEPTIDE 20 96 ACTIVATION PEPTIDE.
 FT CHAIN 97 580 PLASMIN SHORT FORM OF CHAIN A.
 FT CHAIN 581 810 PLASMIN LIGHT CHAIN B.
 FT DOMAIN 103 181 KRINGLE 1.
 FT DOMAIN 184 262 KRINGLE 2.
 FT DOMAIN 275 352 KRINGLE 3.
 FT DOMAIN 377 454 KRINGLE 4.
 FT DOMAIN 481 560 KRINGLE 5.
 FT DOMAIN 581 810 SERINE PROTEASE.
 FT ACT_SITE 622 CHARGE RELAY SYSTEM.
 FT ACT_SITE 665 CHARGE RELAY SYSTEM.
 FT ACT_SITE 760 CHARGE RELAY SYSTEM.
 FT BINDING 136 OMEGA-AMINOCARBOXYLIC ACIDS.

PT BINDING 158 158 OMEGA-AMINOCARBOXYLIC ACIDS.
 PT BINDING 172 172 OMEGA-AMINOCARBOXYLIC ACIDS.
 PT BINDING 432 432 OMEGA-AMINOCARBOXYLIC ACIDS.
 PT BINDING 445 445 OMEGA-AMINOCARBOXYLIC ACIDS.
 PT BINDING 134 134 FIBRIN.
 PT BINDING 136 136 FIBRIN.
 FT DISULFID 49 73 BY SIMILARITY.
 FT DISULFID 53 61 BY SIMILARITY.
 FT DISULFID 103 181 BY SIMILARITY.
 FT DISULFID 124 164 BY SIMILARITY.
 FT DISULFID 152 176 BY SIMILARITY.
 FT DISULFID 185 262 BY SIMILARITY.
 FT DISULFID 188 316 BY SIMILARITY.
 FT DISULFID 206 245 BY SIMILARITY.
 FT DISULFID 234 257 BY SIMILARITY.
 FT DISULFID 275 352 BY SIMILARITY.
 FT DISULFID 296 335 BY SIMILARITY.
 FT DISULFID 324 347 BY SIMILARITY.
 FT DISULFID 377 454 BY SIMILARITY.
 FT DISULFID 398 437 BY SIMILARITY.
 FT DISULFID 426 449 BY SIMILARITY.
 FT DISULFID 481 560 BY SIMILARITY.
 FT DISULFID 502 543 BY SIMILARITY.
 FT DISULFID 531 555 BY SIMILARITY.
 FT DISULFID 567 685 BY SIMILARITY.
 FT DISULFID 577 623 BY SIMILARITY.
 FT DISULFID 607 623 BY SIMILARITY.
 FT DISULFID 699 766 BY SIMILARITY.
 FT DISULFID 729 745 BY SIMILARITY.
 FT DISULFID 756 784 BY SIMILARITY.
 FT CARBOHYD 365 365 O-LINKED (GALNAc...) (BY SIMILARITY).
 SQ SEQUENCE 810 AA; 90255 MW; A75E1C51A1A0F24A CRC64;
 Query Match 26.9%; Score 149; DB 1; Length 810;
 Best Local Similarity 36.4%; Pred. No. 5.4e-09;
 Matches 36; Conservative 12; Mismatches 35; Indels 16; Gaps 6;
 QY 3 CYEGNGHEFYRGKASTDTMGRECLPNSATVLQOYVHAHR---SDALQLGLGKHVCRNPD 59
 Db 377 CHHGQGYRGTSITTTTGKXQSWSS-----MTPHHEKTPENFPNAGL-TMYVCRNPD 430
 QY 60 NRRRPWCYVQVGLKPLV--QECVHDC--DGKXPSPP 94
 Db 431 ADKGWCF--TDPVSRWEYCNLKCSGTEGVAAPP 466
 RESULT 30
 KRM2 HUMAN
 ID KRM2 HUMAN STANDARD; PRT; 462 AA.
 AC Q8NCW0; Q8N2J4; Q8NCW1; Q96GL8; Q9BTP9;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Kremen protein 2 precursor (Kringle-containing protein marking the eye
 DE and the nose) (Dickkopf receptor 2).
 GN KREMEN2 OR KRM2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP Tanaka S., Sugimachi K.;
 RA "Human Kremen2 and Wnt signaling";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
 RC TISSUE=Ovarian carcinoma;
 RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
 RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
 RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
 RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;

RT "NEDO human cDNA sequencing project";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RC MEDLINE=22388257; PubMed=12477932;
RX Klausner R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Strausberg R.L., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H.A., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
to block Wnt/beta-catenin signaling. Forms a ternary complex with
Dkk1 and LRP6 and induces rapid endocytosis and removal of the Wnt
receptor LRP6 from the plasma membrane (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1;
CC IsoId=Q8NCW0-1; Sequences=Displayed;
CC Name=2; Synonyms=Kremen2a;
CC IsoId=Q8NCW0-2; Sequence=VSP_050509, VSP_050510;
CC Name=3; Synonyms=Kremen2b;
CC IsoId=Q8NCW0-3; Sequence=VSP_050511, VSP_050512;
CC Name=4; Synonyms=Kremen2c;
CC IsoId=Q8NCW0-4; Sequence=VSP_050513, VSP_050514;
CC -1- SIMILARITY: Contains 1 CUB domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB086405; BAC00872.1; -
CC EMBL; AB086355; BAC00823.1; -
CC EMBL; AB086356; BAC00824.1; -
CC EMBL; AB086357; BAC00825.1; -
CC EMBL; AK027669; BAB55281.1; -
CC EMBL; AK075033; BAC11365.1; -
CC EMBL; BC003533; AAH03533.1; -
CC EMBL; BC009383; AAH09383.1; -
CC HSP; P00750; IPK2.
CC Genew; HGNC:18797; KREMEN2.
CC InterPro; IPR000859; CUB.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR002889; WSC.
CC Pfam; PF00431; CUB.1.
CC Pfam; PF00051; kringle.1.
CC Pfam; PF01822; WSC.1.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle.1.
CC SMART; SM00042; CUB.1.
CC SMART; SM00130; KR.1.

DR SMART; SM00321; WSC.1.
DR PROSITE; PS01180; CUB.1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Wnt signaling pathway; Glycoprotein; Kringle; Signal; Transmembrane;
KW Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 462
FT DOMAIN 26 364
FT TRANSMEM 365 387
FT DOMAIN 388 462
FT DOMAIN 35 119
FT DOMAIN 121 215
FT DOMAIN 219 326
FT CARBOHYD 49 49
FT CARBOHYD 222 222
FT CARBOHYD 244 244
FT CARBOHYD 351 351
FT VARSPPLIC 394 424
FT VARSPPLIC 425 462
FT VARSPPLIC 367 420
FT VARSPPLIC 421 462
FT VARSPPLIC 367 399
FT VARSPPLIC 400 462
FT CONFLICT 164 202
FT CONFLICT 285 285
FT CONFLICT 462 AA; 48849 MW; CE33015917A9AA68 CRC64;
SQ SEQUENCE 26.8%; Score 148.5; DB 1; Length 462;
Best Local Similarity 42.0%; Pred. No. 3.4e-09;
Matches 29; Conservative 9; Mismatches 26; Indels 5; Gaps 2;
QY 3 CYEKGHFYRG---KASTDTMGPRCLPNSATVLTQTYHAHRSDALQLGKGKHYCNPD 59
DB 36 CFQVNGADYRGHQNRTPRGAGRPCLFWDOTQ--QHSYSSASDPHGRLGGLAHNFCRNP 93
QY 60 NRRRPWCYV 68
DB 94 GDVQPCYV 102
RESULT 31
ID_HGFL_HUMAN STANDARD; PRT; 711 AA.
AC P26927; Q13350; Q14870;
DT 01-AUG-1992 (Rel. 23, Created)
DI 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor-like protein precursor (Macrophage
stimulatory protein) (MSP) (Macrophage stimulating protein).
GN MST1 OR HGFL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92002016; PubMed=1655021;
RA Han S., Stuart L.A., Friesner Degen S.J.;

RT "Characterization of the DNF15S2 locus on human chromosome 3:
 FT identification of a gene coding for four kringle domains with
 RT homology to hepatocyte growth factor.";
 RL Biochemistry 30:9788-9780(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA MEDLINE=93340141; PubMed=8393443;
 RX Yoshimura T., Yuhki N., Wang M.H., Skeel A., Leonard E.J.;
 RT "Cloning, sequencing, and expression of human macrophage stimulating
 protein (MSP, MST1) confirms MSP as a member of the family of kringle
 RT proteins and locates the MSP gene on chromosome 3.";
 RL J. Biol. Chem. 268:15461-15468(1993).
 CC -!- FUNCTION: Probably has no proteolytic activity, since crucial AA
 CC characteristic of serine proteases catalytic sites are not
 CC conserved.
 CC -!- PWM: MAY BE CLEAVED AFTER AA 484, OR YIELD A TWO-CHAIN MOLECULE
 CC HELD TOGETHER BY DISULFIDE BONDS, OR TWO SEPARATE POLYPEPTIDES.
 CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
 CC -!- SIMILARITY: Contains 4 kringle domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; M74178; AAC50165.1; -;
 CC EMBL; U37055; AAC50471.1; -;
 CC EMBL; L11924; AAC59872.1; -;
 CC PIR; A40331; A47136.
 CC HSRP; P00747; 2PK4.
 CC MEROPS; S01.975; -;
 CC Genew; HGNC:7380; MST1.
 CC MIM; 142408; -;
 CC InterPro; IPR009003; Cys Ser trypsin.
 CC InterPro; IPR000001; K-kringle.
 CC InterPro; IPR003014; PAN.
 CC InterPro; IPR003609; Pan_app.
 CC InterPro; IPR001254; Peptidase_S1.
 CC InterPro; IPR001314; Peptidase_S1A.
 CC InterPro; IPR003966; Peptidase_S1A_pr.
 CC Pfam; PF00051; kringle; 4.
 CC Pfam; PF00024; PAN; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC PRINTS; PR01505; PROTHROMBIN.
 CC ProDom; PD000395; Kringle; 4.
 CC SMART; SM00130; KR; 4.
 CC SMART; SM00473; PAN AP; 1.
 CC SMART; SM00020; TRY_SPC; 1.
 CC PROSITE; PS00021; KRINGLE_1; 4.
 CC PROSITE; PS00070; KRINGLE_2; 4.
 CC PROSITE; PS02040; TRYPSIN_DOM; 1.
 CC Kringle; Glycoprotein; Serine protease homolog; Repeat; Signal;
 KW Polymorphism.
 FT SIGNAL 1 18
 FT CHAIN 19 711
 FT DOMAIN 32 109
 FT DOMAIN 110 186
 FT DOMAIN 191 268
 FT DOMAIN 283 361
 FT DOMAIN 370 448
 FT DOMAIN 484 711
 FT DISULFID 56 78
 FT DISULFID 60 66
 FT DISULFID 110 186
 FT DISULFID 131 169
 FT DISULFID 157 181
 FT DISULFID 191 268
 FT DISULFID 191 268

FT DISULFID 194 324
 FT DISULFID 212 251
 FT DISULFID 240 263
 FT DISULFID 283 361
 FT DISULFID 304 343
 FT DISULFID 332 355
 FT DISULFID 370 448
 FT DISULFID 391 431
 FT DISULFID 419 443
 FT DISULFID 468 588
 FT DISULFID 507 523
 FT DISULFID 602 667
 FT DISULFID 632 646
 FT DISULFID 657 685
 FT CARBOHYD 72 72
 FT CARBOHYD 296 296
 FT CARBOHYD 615 615
 FT VARIANT 13 13
 FT VARIANT 212 212
 FT VARIANT 676 676
 FT CONFLICT 623 623
 FT SEQUENCE 711 AA; 80379 MW; 596ED21F180290E4 CRC64;
 SO
 Query Match 26.8%; Score 148.5; DB 1; Length 711;
 Best Local Similarity 33.0%; Pred. No. 5.3e-09;
 Matches 31; Conservative 14; Mismatches 36; Indels 13; Gaps 4;
 QY 2 TCYEGNGHFYRGKASTDTMGKPCLPWNSATVLTQTYHAHRSALQLGLG--KHNYCRNPD 59
 DB 282 SCFRGKGGYRGYGTANTTTAGVPCQRWDA----QIPHQHRFTPEKYACKDLRENFRCNPD 336
 QY 60 NRRRPWCYVQVGLKPLVQEC---WVHDCADGKKP 90
 DB 337 GSEAPWCFT--LRFGMAAFYQIRCTDDVRP 367
 RESULT 32
 ID PLMN_ERIEU STANDARD; PRT; 810 AA.
 AC Q29485;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasminogen precursor (EC 3.4.21.7).
 GN PLG
 OS Brinaceus europaeus (Western European hedgehog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceinae; Brinaceus.
 OX NCBI_TaxID=9365;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RX MEDLINE=96025778; PubMed=7592597;
 RA Lawn R.M., Boonmark N.W., Schwartz K., Lindahl G.E., Wade D.P.,
 RA Byrne C.D., Fong K.J., Meer K., Patthy L.;
 RT "The recurring evolution of lipoprotein(a). Insights from cloning of
 RT hedgehog apolipoprotein(a).";
 RL J. Biol. Chem. 270:24004-24009(1995).
 RN [2]
 RP REVISIONS.
 RA Lawn R.M.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
 CC a proteolytic factor in a variety of other processes including
 CC embryonic development, tissue remodeling, tumor invasion, and
 CC inflammation; in ovulation it weakens the walls of the Graafian
 CC follicle. It activates the urokinase-type plasminogen activator,
 CC collagenases and several complement zymogens, such as C1 and C5.
 CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von
 CC Willebrand factor.

```

-1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
higher selectivity than trypsin. Converts fibrin into soluble
products.
-1- ENZYME REGULATION: Converted into plasmin by plasminogen
activators, both plasminogen and its activator being bound to
fibrin. Cannot be activated with streptokinase.
-1- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
immediately after dissociation from the clot.
-1- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
-1- SIMILARITY: Contains 5 kringle domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL; U33171; AAC48717.1; -.
PIR; I46260; I46260.
HSSP; P00747; IPMK.
MEROPS; S01.233; -.
InterPro; IPR009003; Cys_Ser_trypsin.
InterPro; IPR000011; Kringle.
InterPro; IPR003014; PAN.
InterPro; IPR003609; Pan_app.
InterPro; IPR001254; Peptidase_S1.
InterPro; IPR001314; Peptidase_S1A.
InterPro; IPR003966; Peptidase_S1A_pr.
Pfam; PF00051; Kringle; 5.
Pfam; PF00024; PAN; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; KRINGLE.
PRINTS; PR01505; PROTHROMBIN.
ProDom; PD000395; Kringle; 5.
SMART; SM00130; KR; 5.
SMART; SM00473; PAN AP; 1.
SMART; SM00020; TRYD_SPC; 1.
PROSITE; PS00021; KRINGLE_1; 5.
PROSITE; PS50070; KRINGLE_2; 5.
PROSITE; PS50240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
Signal.
SEQUENCE 1 19 BY SIMILARITY.
CHAIN 20 810 PLASMINOGEN.
CHAIN 20 582 PLASMIN HEAVY CHAIN A (BY SIMILARITY).
CHAIN 583 810 PLASMIN LIGHT CHAIN B (BY SIMILARITY).
DOMAIN 583 810 SERINE PROTEASE.
DOMAIN 183 181 KRINGLE 1.
DOMAIN 185 262 KRINGLE 2.
DOMAIN 275 352 KRINGLE 3.
DOMAIN 379 456 KRINGLE 4.
DOMAIN 482 561 KRINGLE 5.
ACT SITE 622 622 CHARGE RELAY SYSTEM.
ACT SITE 665 665 CHARGE RELAY SYSTEM.
ACT SITE 760 760 CHARGE RELAY SYSTEM.
CARBOHYD 339 339 N-LINKED (GLCNAC...) (POTENTIAL).
SEQUENCE 810 AA; 90902 MW; 5E75780946017A16 CRC64;
Query Match 26.8%; Score 148.5; DB 1; Length 810;
Best Local Similarity 41.5%; Pred. No. 6.1e-09;
Matches 27; Conservative 7; Mismatches 28; Indels 3; Gaps 1;
QY 3 CYEGNGHYRCASDITMGSPCLPWSATVLOQTYHAHSSDALQLGLGHVYCRNPENR 62
379 CYQNGQYRTGTSSTTTGKKCPWTSMRPHRHSKTPENYPADLTM---NYCRNPDGDX 435
DB 63 RPWCY 67

```

```

Query Match      25.8%  Score 148.5;  DB 1;  Length 810;
Best Local Similarity 44.5%;  Pred. No. 6.1e-09;
Matches 27;  Conservative 7;  Mismatches 28;  Indels 3;  Gaps 1;

QY 3  CYEGNGHYFGRKASTDTWGRPCLPWNSATVLOQTYHAHRSDALQIGLKHNYCRNPDNR 62
379  CYQNGQTYRGTSSTTTIGKKCPWTSMRPHRSKTPENYPDADLTM---NYCRNPDGDK 435
QY 63 RPWCY 67

```

PT DISULFID 90 208 INTERCHAIN (BY SIMILARITY).
FT DISULFID 100 108 INTERCHAIN (BY SIMILARITY).
FT DISULFID 130 146 BY SIMILARITY.
FT DISULFID 222 289 BY SIMILARITY.
FT DISULFID 252 268 BY SIMILARITY.
FT DISULFID 279 307 BY SIMILARITY.
FT ACT_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 188 188 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 283 283 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 152 152 STREPTOKINASE-BINDING SITE (PROBABLE).
FT SITE 186 186 STREPTOKINASE-BINDING SITE (PROBABLE).
FT SITE 264 264 STREPTOKINASE-BINDING SITE (PROBABLE).
FT SITE 277 277 SITE OF SUBSTRATE SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 333 AA; 36678 MW; C8C0271B6C6AC8D4 CRC64;
Query Match 26.7%; Score 148; DB 1; Length 333;
Best Local Similarity 38.8%; Pred. No. 2.7e-09;
Matches 40; Conservative 7; Mismatches 40; Indels 16; Gaps 6;
QY 3 CYEGNGHYFGKASTDTMGSPCLPWNATVLQOTYHAHR-----SDALQLGLGKHNYCRNP 58
DB 4 CMFNGKGYGKATTVGIPQCEWAA-----QEPHRSIFTPETNPQAGLEK-NYCRNP 57
QY 59 D-NRRRWCYVQVGLKPLVOECMHDC-----DGKPSPPPEE 96
DB 58 DGDVNGEWCYT-MNQKLFYCDVPQCVCSTSPDCGKQVPEPK 99
RESULT 34
PLMN_PIG STANDARD; PRT; 790 AA.
AC P06867;
DT 01-JAN-1998 (Rel. 06, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen (EC 3.4.21.7).
GN PLG.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RA Schaller J., Marti T., Roessellet S.J., Kaempfer U., Rickli E.E.;
RT "Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the carbohydrate attachment sites with the human and bovine species.";
RL Fibrinolysis 1:91-102(1987).
RN [2]
RP SEQUENCE OF 450-790.
RX MEDLINE=85203907; PubMed=3846533;
RA Marti T., Schaller J., Rickli E.E.;
RT "Determination of the complete amino-acid sequence of porcine miniplasminogen.";
RL Eur. J. Biochem. 149:279-285(1985).
RN [3]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=88185329; PubMed=3356193;
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P., Gerwig G.J., van Halbeek H., Vliegenhart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine plasminogen. Species specificity in relation to sialylation and fucosylation patterns.";
RL Eur. J. Biochem. 173:57-63(1988).
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as a proteolytic factor in a variety of other processes including embryonic development, tissue remodeling, tumor invasion, and inflammation; in ovulation it weakens the walls of the Graafian follicle. It activates the urokinase-type plasminogen activator, collagenases and several complement zymogens, such as C1 and C5. It cleaves fibrin, fibronectin, thrombospondin, laminin and von Willebrand factor.

CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble products
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen activators, both plasminogen and its activator being bound to fibrin. Cannot be activated with streptokinase.
CC -!- PTM: N-LINKED GLYCANS CONTAINS N-ACETYLACTOSAMINE, SIALIC ACID AND IS CORE FUCOSYLATED. O-LINKED GLYCANS CONSIST OF GAL-GALNAc.
CC DISACCHARIDE WITH IS MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).
CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin immediately after dissociation from the clot.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -!- SIMILARITY: Contains 5 kringle domains.
DR PIR: S03733; PLPG.
DR HSSP: P00747; SHPG.
DR MEROPS: S01.233; .
DR GlycoSuiteDB, P06867; .
DR InterPro: IPR009003; Cys_ser_trypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1.
DR InterPro: IPR003966; Peptidase_S1A_dr.
DR Pfam: PF00051; Kringle; 5.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00181; KRINGLE.
DR PRINTS: PR01505; PROTHROMBIN.
DR ProDom: PD000395; Kringle; 5.
DR SMART: SM00130; KR; 5.
DR SMART: SM00473; PAN_AP; 1.
DR SMART: SM00020; TRYF_SPC; 1.
DR PROSITE: PS00021; KRINGLE 1; 5.
DR PROSITE: PS50070; KRINGLE 2; 5.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.
FT CHAIN 1 560 PLASMIN HEAVY CHAIN A.
FT CHAIN 561 790 PLASMIN LIGHT CHAIN B.
FT DOMAIN 561 790 SERINE PROTEASE.
FT DOMAIN 84 162 KRINGLE 1.
FT DOMAIN 166 243 KRINGLE 2.
FT DOMAIN 256 333 KRINGLE 3.
FT DOMAIN 358 435 KRINGLE 4.
FT DOMAIN 461 540 KRINGLE 5.
FT ACT_SITE 602 602 CHARGE RELAY SYSTEM.
FT ACT_SITE 645 645 CHARGE RELAY SYSTEM.
FT ACT_SITE 740 740 CHARGE RELAY SYSTEM.
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .).
FT CARBOHYD 340 340 /FTID=CAR_000019.
FT CARBOHYD 340 340 O-LINKED (GALNAc. . .).
FT /FTID=CAR_000020.
SQ SEQUENCE 790 AA; 88592 MW; F04EA06E74BCD58E CRC64;
Query Match 26.7%; Score 148; DB 1; Length 790;
Best Local Similarity 38.2%; Pred. No. 6.8e-09;
Matches 34; Conservative 8; Mismatches 33; Indels 14; Gaps 5;
QY 3 CYEGNGHYFGKASTDTMGSPCLPWNATVLQOTYHAHR-----SDALQLGLGKHNYCRNP 59
DB 358 CYRNGESYRGTSSTITGRKQSW-----VSMTPHREKTPGFPNAGL-TMYCRNP 411
QY 60 NERRWCYVQVGLKPLV--QECMHDCAD 86
DB 412 ADKSPWCYT---TDPRWYCNLKKCSE 437
RESULT 35

HGFL_MOUSE
ID HGFL_MOUSE STANDARD; PRT; 716 AA.
AC P26928;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hepatocyte growth factor-like protein precursor (Macrophage
stimulatory protein) (MSP).
GN MST1 OR HGFL.
OS Mus musculus. (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Liver;
RX MEDLINE=92002017; PubMed=1832957;
RT Friesner Degen S.G.; Stuart L.A.; Han S.; Jamison C.S.;
RT "Characterization of the mouse cDNA and gene coding for a hepatocyte
growth factor-like protein: expression during development.";
RL Biochemistry 30:9781-9791(1991).
CC -!- FUNCTION: Probably has no proteolytic activity, since crucial AA
characteristic of serine proteases catalytic sites are not
conserved.
CC -!- TISSUE SPECIFICITY: Liver. Lower levels in lung, placenta and
adrenal.
CC -!- DEVELOPMENTAL STAGE: Is expressed at low levels during gestation.
CC Just before birth the level increases dramatically and remains
stable afterwards.
CC -!- PTM: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE
HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE
POLYPEPTIDES.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -!- SIMILARITY: Contains 4 kringle domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC ENBL; M74180; AAA50166.1; -.
CC ENBL; M74181; AAA50167.1; -.
CC PIR; A40332; A40332.
CC HSP; P00747; IGRN.
CC MEROPS; S01.975; -.
CC MGD; MGI:96080; Mst1.
CC GO; GO:0007566; P:embryo implantation; IC.
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR003014; PAN.
CC InterPro; IPR003609; PAN_app.
CC InterPro; IPR001254; Peptidase S1.
CC InterPro; IPR001314; Peptidase S1A.
CC InterPro; IPR003966; Peptidase_S1A_pr.
CC Pfam; PF00051; kringle; 4.
CC Pfam; PF00024; PAN; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC PRINTS; PR01505; PROTHROMBIN.
CC ProDom; PD000395; Kringle; 4.
CC SMART; SM00130; KR; 4.
CC SMART; SM00473; PAN AP; 1.
CC SMART; SM00020; TRYPSIN; 1.
CC PROSITE; PS00021; KRINGLE_1; 4.
CC PROSITE; PS00070; KRINGLE_2; 4.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
KW Kringle; Glycoprotein; Serine protease homolog; Repeat; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 716 HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.

FT DOMAIN 19 109 PAP.
FT DOMAIN 110 186 KRINGLE 1.
FT DOMAIN 191 268 KRINGLE 2.
FT DOMAIN 292 370 KRINGLE 3.
FT DOMAIN 379 457 KRINGLE 4.
FT DOMAIN 489 716 SERINE PROTEASE-LIKE.
FT DISULFID 56 78 BY SIMILARITY.
FT DISULFID 60 66 BY SIMILARITY.
FT DISULFID 110 186 BY SIMILARITY.
FT DISULFID 131 169 BY SIMILARITY.
FT DISULFID 157 181 BY SIMILARITY.
FT DISULFID 191 268 BY SIMILARITY.
FT DISULFID 194 333 INTERCHAIN (BY SIMILARITY).
FT DISULFID 212 251 BY SIMILARITY.
FT DISULFID 240 263 BY SIMILARITY.
FT DISULFID 292 370 BY SIMILARITY.
FT DISULFID 313 352 BY SIMILARITY.
FT DISULFID 341 364 BY SIMILARITY.
FT DISULFID 379 457 BY SIMILARITY.
FT DISULFID 400 440 BY SIMILARITY.
FT DISULFID 428 452 BY SIMILARITY.
FT DISULFID 477 593 INTERCHAIN (BY SIMILARITY).
FT DISULFID 512 528 BY SIMILARITY.
FT DISULFID 607 672 BY SIMILARITY.
FT DISULFID 637 651 BY SIMILARITY.
FT DISULFID 662 690 BY SIMILARITY.
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 19 19 P -> O (IN GENOMIC SEQUENCE).
SQ SEQUENCE 716 AA; 80588 MW; BBCE02EF85213ACC CRC64;
Query Match 26.6%; Score 147.5; DB 1; Length 716;
Best Local Similarity 36.5%; Pred. No. 7e-09;
Matches 38; Conservative 11; Mismatches 34; Indels 21; Gaps 7;
Qy 3 CYENGHFYKASTDTMGRCPLPWSATVLTQYHAHR-----SDALQLGKHNYCRN 57
Db 379 CYHGGQYRGSRVSKTKGVQCQHSSET-----PHKQFTPTSPAQGL-EANFCRN 430
Qy 58 PD-NRRPWCYVQVGLKP--LVQECMVHDCADGKKPS--SPPEE 96
Db 431 PDGDSHGWCYT---LDPDLFDYCALQRCDDQPPSILDPDQ 471
RESULT 36
ID ROR2 HUMAN STANDARD; PRT; 943 AA.
AC Q01974; Q9HAV7; Q9HB61;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor ROR2 precursor
DE (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 2).
GN ROR2 OR NTRKR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93100347; PubMed=1334494;
RA Maslakowski P.; Carroll R.D.;
RT "A novel family of cell surface receptors with tyrosine kinase-like
domain.";
RL J. Biol. Chem. 267:26181-26190(1992).
RN [2]
RP SEQUENCE OF 34-943 FROM N.A., AND VARIANT ILE-819.
RX MEDLINE=20164326; PubMed=10700182;
RA Oldridge M.; Fortuna A.M.; Maringa M.; Propping P.; Mansour S.;
RA Pollitt C.; DeChiara T.M.; Kimble R.B.; Valenzuela D.M.;
RA Yancopoulos G.D.; Wilkie A.O.M.;

"Dominant mutations in ROR2, encoding an orphan receptor tyrosine kinase, cause brachydactyly type B";
 Nat. Genet. 24:275-278(2000).
 [3]
 RP SEQUENCE OF 34-574 FROM N.A. AND VARIANT THR-245.
 RX MEDLINE=20442029; PubMed=10986040;
 RA Schwabe G.C., Tinschert S., Buschow C., Weinsche P., Wolff G.,
 RA Gillesen-Kaeschach G., Oldridge M., Wilkie A.O.M., Koeneke R.,
 RA Mundlos S.,
 RT "Distinct mutations in the receptor tyrosine kinase gene ROR2 cause
 brachydactyly type B";
 RL Am. J. Hum. Genet. 67:822-831(2000).
 RN [4]
 RP VARIANTS RRS CYS-184; TRP-189; TRP-366 AND LYS-620.
 RX MEDLINE=20392394; PubMed=10932186;
 RA Afzal A.R., Rajab A., Fenske C.D., Oldridge M., Elanko N.,
 RA Ternes-Pereira E., Tucsnyez B., Munday V.A., Patton M.A.,
 RA Wilkie A.O.M., Jeffery S.,
 RT "Recessive Robinow syndrome, allelic to dominant brachydactyly type B,
 is caused by mutation of ROR2";
 RL Nat. Genet. 25:419-422(2000).
 RN [5]
 RP VARIANT RRS TYR-182.
 RX MEDLINE=20392395; PubMed=10932187;
 RA van Bokhoven H., Celli J., Kayserili H., van Beusekom E., Balci S.,
 RA Brussel W., Skovby F., Kerr B., Percin E.F., Akarsu N., Brunner H.G.,
 RL Nat. Genet. 26:383-383(2000).
 CC -!- FUNCTION: Tyrosine-protein kinase receptor which may be involved
 in the early formation of the chondrocytes. It seems to be
 required for cartilage and growth plate development.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DEVELOPMENTAL SPAGS: Expressed at high levels during early
 embryonic development. The expression levels drop strongly around
 day 16 and there are only very low levels in adult tissues.
 CC -!- DISEASE: Defects in ROR2 are a cause of brachydactyly type B1
 (BDB1) [MIM:113000]. BDB1 is an autosomal dominant skeletal
 disorder characterized by hypoplasia/aplasia of distal phalanges
 and nails. In BDB1 the middle phalanges are short but in addition
 the terminal phalanges are rudimentary or absent. Both fingers and
 toes are affected. The thumbs and big toes are usually deformed.
 CC -!- DISEASE: Defects in ROR2 are a cause of recessive Robinow syndrome
 (RRS) [MIM:268310]. RRS is an autosomal disorder characterized by
 skeletal dysplasia with generalized limb bone shortening,
 segmental defects of the spine, brachydactyly and a dysmorphic
 facial appearance.
 CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR
 subfamily.
 CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M97639; AAA60276.1; -;
 DR EMBL; AF294796; AAG01184.2; -;
 DR EMBL; AF254747; AAG01184.2; JOINED.
 DR EMBL; AF254748; AAG01184.2; JOINED.
 DR EMBL; AF254749; AAG01184.2; JOINED.

DR EMBL; AF254750; AAG01184.2; JOINED.
 DR EMBL; AF254751; AAG01184.2; JOINED.
 DR EMBL; AF254752; AAG01184.2; JOINED.
 DR EMBL; AF254753; AAG01184.2; JOINED.
 DR EMBL; AF279762; AAG33132.1; -;
 DR EMBL; AF279765; AAG33132.1; JOINED.
 DR EMBL; AF279766; AAG33132.1; JOINED.
 DR EMBL; AF279767; AAG33132.1; JOINED.
 DR EMBL; AF279768; AAG33132.1; JOINED.
 DR EMBL; AF279769; AAG33132.1; JOINED.
 DR EMBL; AF279770; AAG33132.1; JOINED.
 DR EMBL; AF279771; AAG33132.1; JOINED.
 DR PIR; B45082; B45082.
 DR HSPF; P00747; LKRN.
 DR Genew; HGNC:10257; ROR2.
 DR MIM; 602337; -;
 DR MIM; 113000; -;
 DR MIM; 268310; -;
 DR GO; GO:0005887; C: integral to plasma membrane; TAS.
 DR GO; GO:0004714; F: transmembrane receptor protein tyrosine kin. . . ; TAS.
 DR GO; GO:0007275; P: development; TAS.
 DR GO; GO:0007185; P: signal transduction; TAS.
 DR InterPro; IPR000024; Fz domain.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG C2.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR InterPro; IPR008266; Tyr_kinase_AS.
 DR Pfam; PF01392; Fz; 1.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000395; Kringle; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00219; Tyrc; 1.
 DR PROSITE; PS00038; FZ; 1.
 DR PROSITE; PS00835; IG LIKE; 1.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 1.
 DR PROSITE; PS00107; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00011; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Receptor;
 KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
 KW Immunoglobulin domain; Developmental protein; Polymorphism;
 KW Disease mutation.
 FT SIGNAL 1 33
 FT CHAIN 34 943
 FT TYROSINE-PROTEIN KINASE TRANSMEMBRANE
 FT RECEPTOR ROR2.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE C2-TYPE.
 FT FZ.
 FT KRINGLE.
 FT PROTEIN KINASE.
 FT SER/THR-RICH.
 FT PRO-RICH.
 FT SER/THR-RICH.
 FT ATP (BY SIMILARITY).
 FT BINDING 507 507
 FT ACT SITE 615 615
 FT MOD_RES 646 646
 FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT BY SIMILARITY.
 FT N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT C -> Y (in RRS).
 FT VARIANT 182 182

between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL; AB010384; BAA75481.1; --
HSPG; P00747; SHFG.
MGI; 1347521; Ror2.
GO; GO:0001501; P:skeletal development; IMP.
InterPro; IPR000024; Fz domain.
InterPro; IPR007110; IG-like.
InterPro; IPR003598; IG_C2.
InterPro; IPR000001; Kringle.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR001245; Tyr_kinase.
InterPro; IPR008366; Tyr_kinase_AS.
Pfam; PF01392; Fz; 1.
Pfam; PF00047; IG; 1.
Pfam; PF00051; kringle; 1.
Pfam; PF00069; pkinase; 1.
PRINTS; PR00018; KRINGLE.
PRINTS; PR00109; TYRKINASE.
ProDom; PD000395; Kringle; 1.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00408; IGC2; 1.
SMART; SM00130; KR; 1.
SMART; SM00219; TyrcK; 1.
PROSITE; PS00038; FZ; 1.
PROSITE; PS00835; IG LIKE; 1.
PROSITE; PS00021; KRINGLE 1; 1.
PROSITE; PS00070; KRINGLE 2; 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
Transfaser; tyrosine-protein kinase; ATP-binding; Receptor;
Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
Immunoglobulin domain; Developmental protein.
SIGNAL 1 33 POTENTIAL.
CHAIN 34 944 TYROSINE-PROTEIN KINASE TRANSMEMBRANE
RECEPTOR ROR2
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
IG-LIKE C2-TYPE.
FZ.
KRINGLE.
PROTEIN KINASE.
SER/THR-RICH.
PRO-RICH.
SER/THR-RICH.
ATP (BY SIMILARITY).
ATP (BY SIMILARITY).
BY SIMILARITY.
PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
BY SIMILARITY.
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
SEQUENCE 944 AA; 105050 MW; CD2EEBC710387A56 CRC64;

Query Match 26.6%; Score 147.5; DB 1; Length 944;
Best Local Similarity 41.4%; Pred. No. 9.3e-09;
Matches 36; Conservative 9; Mismatches 31; Indels 11; Gaps 6

QY 3 CYEGNGHFVRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGHKNYCRPNDR 59
Db 316 CYNGSGADYRGMASTTKSGHQCPW---ALQHP-HSHLSSTFPELG-GGHAYCRNPGQM 369

QY 60 NRRR-PWCYVQVGLKPLVQECMVHDC 85
Db 370 GOMEQWCFQ-QKNVRLVCLDVPCS 395

This SWISS-PROT entry is copyright. It is produced through a collaboration

```

Query Match      26.4%; Score 146.5; DB 1; Length 461;
Best Local Similarity 42.0%; Pred. No. 5.7e-09;
Matches 29; Conservative 8; Mismatches 27; Indels 5; Gaps 2;

QY 3 CYEGNGHFYRGK---ASTDTMGPRCLPWSATVLTQTYHAHRSDALQLGLGKKNYCRNPD 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35 CFVNGADYRGHQNYTGPRGAGRPCLFWDQDQ--OHSYSSASDPQGRWGLGAHNFRCRNP 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 60 NRRRPWCYV 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 93 GDVQPCYV 101

RESULT 39
PLMN HUMAN
ID PLMN HUMAN STANDARD; PRT; 810 AA.
AC P00747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT ASN-472.
RX MEDLINE=90203879; PubMed=2318848;
RA Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
   in the fibrinolytic system.";
RL J. Biol. Chem. 265:6104-6111(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=3030813;
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
   for human plasminogen.";
RL FEBS Lett. 213:254-260(1987).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANTS LYS-57; GLN-133; HIS-261; TRP-408;
RP ASN-472; VAL-494 AND TRP-523.
RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 20-810, AND VARIANT ASN-472.
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the PIR data bank.
RN [5]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
   human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
RN [6]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
   of human plasminogen and their interaction with the NH2-terminal
   activation peptide as studied by affinity chromatography.";
RL Eur. J. Biochem. 50:489-494(1975).
RN [7]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND VARIANT ASN-472.
RA Sottrup-Jensen L., Glaesys H., Zajdel M., Petersen T.E., Magnusson S.;
RL (in) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);
   Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
   Raven Press, New York (1978).
RN [8]
RP SEQUENCE OF 483-604.

```

RX MEDLINE=76043692; PubMed=126863;
 RA Wiman B., Wallen P.;
 RT "Amino-acid sequence of the cyanogen-bromide fragment from human
 plaminogen that forms the linkage between the plasmin chains.";
 RL Eur. J. Biochem. 58:539-547(1975).
 RN [9]
 RP SEQUENCE OF 581-810.
 RX MEDLINE=77225245; PubMed=142009;
 RA Wiman B.;
 RT "Primary structure of the B-chain of human plasmin.";
 RL Eur. J. Biochem. 76:129-137(1977).
 RN [10]
 RP ACTIVE SITE.
 RX MEDLINE=73149248; PubMed=4594729;
 RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
 RT "The primary structure of human plasminogen. II. The histidine loop
 of human plasmin: light (B) chain active center histidine sequence.";
 RL J. Biol. Chem. 248:1631-1633(1973).
 RN [11]
 RP ACTIVE SITE.
 RX MEDLINE=69234739; PubMed=4240117;
 RA Grosekopf W.R., Summaria L., Robbins K.C.;
 RT "Studies on the active center of human plasmin. Partial amino acid
 sequence of a peptide containing the active center serine residue.";
 RL J. Biol. Chem. 244:3590-3597(1969).
 RN [12]
 RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
 RX MEDLINE=82213905; PubMed=6919539;
 RA Trexler M., Valli Z., Pathy L.;
 RT "Structure of the omega-aminocarboxylic acid-binding sites of human
 plasminogen. Arginine 70 and aspartic acid 56 are essential for
 binding of ligand by kringles 4.";
 RL J. Biol. Chem. 257:7401-7406(1982).
 RN [13]
 RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
 RX MEDLINE=85054794; PubMed=6094526;
 RA Valli Z., Pathy L.;
 RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
 are essential for fibrin affinity of the kringles 1 domain.";
 RL J. Biol. Chem. 259:13690-13694(1984).
 RN [14]
 RP PHOSPHORYLATION SITE SER-597.
 RX MEDLINE=97345939; PubMed=9201958;
 RA Wang H., Prokoc M., Bretthauer R.K., Castellino F.J.;
 RT "Serine-578 is a major phosphorylation locus in human plasma
 plasminogen.";
 RL Biochemistry 36:8100-8106(1997).
 RN [15]
 RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=88185329; PubMed=3356193;
 RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
 Gerwig G.J., van Halbeek H., Vliegthart J.F.;
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
 plasminogen. Species specificity in relation to sialylation and
 fucosylation patterns.";
 RL Eur. J. Biochem. 173:57-63(1988).
 RN [16]
 RP CARBOHYDRATE-LINKAGE SITE SER-268.
 RX MEDLINE=97207306; PubMed=9054441;
 RA Pirie-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
 Pizzo S.V.;
 RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of
 human plasminogen 2.";
 RL J. Biol. Chem. 272:7408-7411(1997).
 RN [17]
 RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
 RX MEDLINE=95042728; PubMed=7525077;
 RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
 Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
 RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
 suppression of metastases by a Lewis lung carcinoma.";
 RL Cell 79:315-328(1994).
 RN [18]

RP CHARACTERIZATION OF ANGIOSTATIN.
 RX MEDLINE=97238710; PubMed=9102221;
 RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,
 Lapcevic R., Nacy C.A.;
 RT "A recombinant human angiotensin protein inhibits experimental primary
 and metastatic cancer.";
 RL Cancer Res. 57:1329-1334(1997).
 RN [19]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
 RX MEDLINE=92031502; PubMed=1657148;
 RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
 RT "Crystal and molecular structure of human plasminogen kringles 4
 refined at 1.9-A resolution.";
 RL Biochemistry 30:10576-10588(1991).
 RN [20]
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
 RX MEDLINE=92031503; PubMed=1657149;
 RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
 RT "The refined structure of the epsilon-aminocaproic acid complex of
 human plasminogen kringles 4.";
 RL Biochemistry 30:10589-10594(1991).
 RN [21]
 RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
 RX MEDLINE=96180681; PubMed=8611560;
 RA Mathews I.I., Vanderhoff-Hanover P., Castellino F.J., Tulinsky A.;
 RT "Crystal structures of the recombinant kringles 1 domain of human
 plasminogen in complexes with the ligands epsilon-aminocaproic acid
 and trans-4-(aminomethyl)cyclohexane-1-carboxylic acid.";
 RL Biochemistry 35:2567-2576(1996).
 RN [23]
 RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
 RX MEDLINE=98198034; PubMed=9521645;
 RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
 Castellino F.J.;
 RT "Structure and ligand binding determinants of the recombinant kringles
 5 domain of human plasminogen.";
 RL Biochemistry 37:3258-3271(1998).
 RN [24]
 RP STRUCTURE BY NMR OF 96-184.
 RX MEDLINE=94237157; PubMed=8181475;
 RA Rejzante M.R., Llinas M.;
 RT "1H-NMR assignments and secondary structure of human plasminogen
 kringles 1.";
 RL Eur. J. Biochem. 221:927-937(1994).
 RN [25]
 RP STRUCTURE BY NMR OF 96-184.
 RX MEDLINE=94237158; PubMed=8181476;
 RA Rejzante M.R., Llinas M.;
 RT "Solution structure of the epsilon-aminohexanoic acid complex of
 human plasminogen kringles 1.";
 RL Eur. J. Biochem. 221:939-949(1994).
 RN [26]
 RP STRUCTURE BY NMR OF 183-354.
 RX MEDLINE=96194156; PubMed=8652577;
 RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
 Rickli E.E.;
 RT "Recombinant gene expression and 1H NMR characteristics of the
 kringles (2 + 3) supermodule: spectroscopic/functional individuality
 of plasminogen kringles domains.";
 RL Biochemistry 35:2357-2364(1996).
 RN [27]
 RP STRUCTURE BY NMR OF 374-461.
 RX MEDLINE=90219023; PubMed=2157850;

Query Match 26.4%; Score 146.5; DB 1; Length 810;
 Best Local Similarity 37.4%; Pred. No. 1e-08;
 Matches 37; Conservative 13; Mismatches 32; Indels 17; Gaps 7;

QY 3 CYBGNHGFYRGKASTDMGRPCLPWNASATVLOQTYHAHR---SDALQLGLGKHNYCRNPD 59
DB 377 CYHGDGQSYRGTSSTTTGKKQCSWSS-----MTPHRHQKTPENYPNAGL-TMNYCRNPD 430
QY 60 NRPRPCVYQVGLKPLV--QECMVHDCADGKKPS--SPP 94
DB 431 ADKGPWCFT---TDPSSVRWEYCNLKKCS-GTEASVVAPP 465

RESULT 40
ID -ROR1_HUMAN STANDARD; PRT; 937 AA.
AC Q01973; Q92776;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor ROR1 precursor
DE (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 1).
GN ROR1 OR NTRK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN J. Biol. Chem. 267:26181-26190(1992).
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=93100347; PubMed=1334494;
RA Masiakowski P., Carroll R.D.;
RT "A novel family of cell surface receptors with tyrosine kinase-like domain.";
RT J. Biol. Chem. 267:26181-26190(1992).
RN (2)
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RX MEDLINE=97030043; PubMed=8875995;
RA Reddy U.R., Phatak S., Pleasure D.;
RT "Human neural tissues express a truncated Ror1 receptor tyrosine kinase, lacking both extracellular and transmembrane domains.";
RL Oncogene 13:1555-1559(1996).
CC -!- FUNCTION: Tyrosine-protein kinase receptor whose role is not yet clear.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q01973-1; Sequence=Displayed;
CC Name=Short; Synonyms=T-ROR1;
CC IsoId=Q01973-2; Sequence=VSP_005008;
CC -!- TISSUE SPECIFICITY: Expressed strongly in human heart, lung, and kidney, but weakly in the CNS. The short isoform is strongly expressed in fetal and adult CNS and in a variety of human cancers, including those originating from CNS or PNS neuroectoderm.
CC -!- DEVELOPMENTAL STAGE: Expressed at high levels during early embryonic development. The expression levels drop strongly around day 16 and there are only very low levels in adult tissues.
CC -!- SIMILARITY: Belongs to the tyr family of protein kinases. ROR subfamily.
CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M97675; AAC60275.1; -;
CC EMBL; U38894; AAC50714.1; -;

PIR; A45082; A45082.
HSSP; P00747; ICEA.
Genew; HGNC:10256; ROR1.
MIM; 602336; -;
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0005987; C:integral to plasma membrane; TAS.
DR GO; GO:0004714; F:transmembrane receptor protein tyrosine kin. ; TAS.
DR GO; GO:0007169; F:transmembrane receptor protein tyrosine kin. ; TAS.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig c2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00219; Ty-Kc; 1.
DR PROSITE; PS00038; Fz; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00111; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_DOM; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Receptor;
KW Transmembrane; signal; Glycoprotein; Kringle; Phosphorylation;
KW Immunoglobulin domain; Alternative splicing.
FT SIGNAL 1 29
FT CHAIN 30 937
FT TYROSINE-PROTEIN KINASE TRANSMEMBRANE
FT RECEPTOR ROR1.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT IG-LIKE C2-TYPE.
FT FZ.
FT KRINGLE.
FT PROTEIN KINASE.
FT SER/THR-RICH.
FT PRO-RICH.
FT SER/THR-RICH.
FT ATP (BY SIMILARITY).
FT ATP (BY SIMILARITY).
FT BY SIMILARITY.
FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT BY SIMILARITY.
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT Missing (in isoform Short).
FT /FTid=VSP_005008.
SQ SEQUENCE 937 AA; 104312 MW; 0D0694DBF29F4773 CRC64;
Query Match 26.4%; Score 146; DB 1; Length 937;
Best Local Similarity 44.1%; Pred. No. 1.4e-08;
Matches 30; Conservative 5; Mismatches 25; Indels 8; Gaps 3;
QY 3 CYBGNHGFYRGKASTDMGRPCLPWNASATVLOQTYHAHRSDALQLGLGKHNYCRNPD 60
DB 313 CYNSTGVYRGVTSVTKSGRCQWPNS-----QYPHTHTFTALFPFLNGHSGYCRNPGN 367
QY 61 RRR-PWCY 67
DB 368 QKEAPWCF 375

Transferase; Tyrosine-protein kinase; ATP-binding; Receptor;
Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
Immunoglobulin domain.
KW SIGNAL 1 29 POTENTIAL
KW CHAIN 30 937 RECEPTOR ROR1.
FT DOMAIN 30 406 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 407 427 POTENTIAL.
FT DOMAIN 428 937 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 42 141 IG-LIKE C2-TYPE.
FT DOMAIN 165 299 FZ.
FT DOMAIN 312 391 KRINGLE.
FT DOMAIN 473 746 PROTEIN KINASE.
FT DOMAIN 753 782 SER/THR-RICH.
FT DOMAIN 784 851 PRO-RICH.
FT DOMAIN 853 876 SER/THR-RICH.
FT NP_BIND 479 487 ATP (BY SIMILARITY).
FT BINDING 506 506 ATP (BY SIMILARITY).
FT ACT_SITE 615 615 BY SIMILARITY.
FT MOD_RES 645 645 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT DISULFID 79 131 BY SIMILARITY.
FT CARBOHYD 47 47 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 937 AA; 104156 MW; D728733E67D1782C CRC64;
Query Match 26.4%; Score 146; DB 1; Length 937;
Best Local Similarity 44.1%; Pred. No. 1.4e-08;
Matches 30; Conservative 5; Mismatches 25; Indels 8; Gaps 3;
QY 3 CYEENGHFGKASTDTMGPRCLPWNSATVLQYTHAHRSDALQLG--LGSHNYCRNPDN 60
Db 313 CYNSTGVDRYGVTSVTKSGRCQCPWNS-----QYPHTSFALRFPPLNGHSHYCRNPGN 367
QY 61 RRR-PWCY 67
Db 368 QKEAPWCF 375
RESULT 42
PLMN_MOUSE STANDARD; PRT; 812 AA.
ID PLMN_MOUSE STANDARD; PRT; 812 AA.
AC P20918; Q8CIS2; Q91WJ5;
DT 01-FEB-1991 (Rel. 17, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91184812; PubMed=2081600;
RA Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;
RT "Characterization of the cDNA coding for mouse plasminogen and
RT localization of the gene to mouse chromosome 17.";
RL Genomics 8:49-61(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RA Brathwaite M., Maelzt P., Qian Y., Dudekula D., Schlessinger D.,
RA Nagaraja R.;
RT "Genomic sequence analysis in the mouse t-complex region.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RESULT 41
ROR1_MOUSE STANDARD; PRT; 937 AA.
ID ROR1_MOUSE STANDARD; PRT; 937 AA.
AC Q9Z139;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor ROR1 precursor
DE [EC 2.7.1.112] (Neurotrophic tyrosine kinase, receptor-related 1)
DE (ROR1).
GN ROR1 OR NTRK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99248426; PubMed=102311392;
RA Oishi I., Takeuchi S., Hashimoto R., Nagabukuro A., Ueda T., Liu Z.J.,
RA Hatta T., Akira S., Matsuda Y., Yamamura H., Otani H., Minami Y.;
RT "Spatio-temporally regulated expression of receptor tyrosine kinases,
RT mRor1, mRor2, during mouse development: implications in development
RT and function of the nervous system.";
RL Genes Cells 4:41-56(1999).
CC -!- FUNCTION: Tyrosine-protein kinase receptor whose role is not yet
CC clear.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR
CC subfamily.
CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB010383; BAA75480.1; .
DR HSP; P00747; ICEA
DR MGD; MGI:1347520; Ror1.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TVR; 1.

EMBL; AY134430; AAN15905.1; -	1	19	PLASMINOGEN.
PIR; A38514; PLMS	20	812	PLASMIN HEAVY CHAIN A.
HSSP; P00747; IPMK.	20	581	ACTIVATION PEPTIDE
MEROPS; S01.233; -	20	97	PLASMIN SHORT FORM OF CHAIN A.
MGP; MG1.97620; Pig.	98	581	ANGIOSTATIN.
GO; GO:0016506; F:apoptosis activator activity; IDA.	582	812	PLASMIN LIGHT CHAIN B.
GO; GO:0006916; F:apoptosis; IDA.	103	181	KRINGLE 1.
GO; GO:0009003; Cys Ser trypsin.	184	262	KRINGLE 2.
InterPro; IPR009003; Cys Ser trypsin.	275	352	KRINGLE 3.
InterPro; IPR000001; Kringle.	377	454	KRINGLE 4.
InterPro; IPR003014; PAN.	481	560	KRINGLE 5.
InterPro; IPR003609; Pan_app.	582	812	SERINE PROTEASE.
InterPro; IPR001254; Peptidase_S1.	582	812	CHARGE RELAY SYSTEM (BY SIMILARITY).
InterPro; IPR001314; Peptidase_S1A.	624	624	CHARGE RELAY SYSTEM (BY SIMILARITY).
InterPro; IPR003966; Peptidase_S1A_pr.	667	667	CHARGE RELAY SYSTEM (BY SIMILARITY).
Pfam; PF00051; kringle; 5.	762	762	BY SIMILARITY.
Pfam; PF00024; PAN; 1.	49	73	BY SIMILARITY.
Pfam; PF00089; trypsin; 1.	53	61	BY SIMILARITY.
PRINTS; PR00722; CHYMOTRYPSIN.	103	181	BY SIMILARITY.
PRINTS; PR00018; KRINGLE	124	164	BY SIMILARITY.
PRINTS; PR00105; PROTHROMBIN.	152	176	BY SIMILARITY.
ProDom; PD000395; Kring1; 4.	185	262	BY SIMILARITY.
SMART; SM00130; KR; 5.	188	316	BY SIMILARITY.
SMART; SM00473; PAN AP; 1.	206	245	BY SIMILARITY.
SMART; SM00020; TRY1_SPC; 1.	234	257	BY SIMILARITY.
PROSITE; PS00021; KRINGLE_1; 5.	275	352	BY SIMILARITY.
PROSITE; PS50070; KRINGLE_2; 5.	296	335	BY SIMILARITY.
PROSITE; PS50240; TRYPSIN_DOM; 1.	324	347	BY SIMILARITY.
PROSITE; PS00134; TRYPSIN_HIS; 1.	377	454	BY SIMILARITY.
PROSITE; PS00135; TRYPSIN_SER; 1.	398	437	BY SIMILARITY.
Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;	426	449	BY SIMILARITY.
Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;	481	560	BY SIMILARITY.
Signal.	502	543	BY SIMILARITY.
FT CHAIN	531	555	BY SIMILARITY.
FT CHAIN	568	687	INTERCHAIN (BY SIMILARITY).
FT CHAIN	578	586	INTERCHAIN (BY SIMILARITY).
FT CHAIN	609	625	BY SIMILARITY.
FT CHAIN	701	768	BY SIMILARITY.
FT CHAIN	731	747	BY SIMILARITY.
FT CHAIN	758	786	BY SIMILARITY.
FT CHAIN	235	235	R -> H (IN REF. 1).
FT CHAIN	235	235	CONFLICT

```

FT CONFLICT 525 525 G -> D (IN REF. 1).
FT CONFLICT 649 649 S -> L (IN REF. 1).
SQ SEQUENCE 812 AA; 90781 MW; 2417326056A2FFD2 CRC64;

Query Match 26.1%; Score 144.5; DB 1; Length 812;
Best Local Similarity 32.7%; Pred. No. 1.7e-08;
Matches 33; Conservative 8; Mismatches 47; Indels 13; Gaps 4;

QY 3 CYEGNGHYRGASDTWGRCPPLNWSATVLOOTVHARSALQLGLG--KHNYCRNPDN 60
DB 275 CLKRGNGYRGVTVVSGTKTCQWRS-----EQPHRHNTPEFPCKNLEENYCRNPDG 329

QY 61 RRRPCCYVQVGLKPLVQECMVHCDAGKKP-----SSPPEE 96
DB 330 ETAPWCYT-TDSQLRWEYCEIFSCSSASPDQSDSSVPPEE 369

RESULT 43
RORI DROME
ID RORI DROME STANDARD; PRT; 685 AA.
AC Q24488;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor Ror precursor
DE (EC 2.7.1.112) (GPro).
GN ROR OR CG4926.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=Canton-S; TISSUE=Larval brain;
RX MEDLINE=93348222; PubMed=8394009;
RA Wilson C. Goherdan D.C.I. Steller H.;
RT "Ror", a potential neurotrophic receptor gene, encodes a Drosophila
RT homolog of the vertebrate Ror family of Trk-related receptor tyrosine
RT kinases."
RL Proc. Natl. Acad. Sci. U.S.A. 90:7109-7113(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Change M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.P., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoston K., Houston K.A., Howland T.J., Wei M.-H., Ibegawa C.,
RA Jastoli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```

```

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapietson M., Strong R., Sun E., Wang X.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.N., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE OF 545-597 FROM N.A.
RX MEDLINE=98401146; PubMed=9731193;
RA Oates A.C., Wollberg P., Achen M.G., Wilks A.F.;
RT "Sampling the genomic pool of protein tyrosine kinase genes using the
RT polymerase chain reaction with genomic DNA."
RL Biochem. Biophys. Res. Commun. 249:660-667(1998).
CC -/- FUNCTION: Tyrosine-protein kinase receptor that functions during
CC early stages of neuronal development.
CC -/- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -/- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -/- TISSUE SPECIFICITY: Expressed in neurons of the developing nervous
CC system.
CC -/- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR
CC subfamily.
CC -/- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -/- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L20297; AAA28860.1; -.
DR EMBL; AE003628; AAF52885.1; -.
DR EMBL; AJ002908; CAA05743.1; -.
DR PIR; A48289; A48289.
DR HSP; P11362; IFKG.
DR FlyBase; FBgn0010407; Ror.
DR GO; GO:0016021; C:integral to membrane; NAS.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; NAS.
DR GO; GO:0007417; P:central nervous system development; IEP.
DR GO; GO:0006468; P:protein amino acid phosphorylation; NAS.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002011; Recepttyrkinl.
DR InterPro; IPR001245; Tyr_kinase_AS.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00330; KR; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00038; FZ; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
DR Transmembrane; Signal; Tyrosine-protein kinase; ATP-binding; Receptor;
DR Developmental protein.
FT SIGNAL 1 24
FT CHAIN 25 685 TYROSINE-PROTEIN KINASE TRANSMEMBRANE

```

```

FT DOMAIN 25 317 RECEPTOR_ROR
FT TRANSNM 318 338 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 339 685 POTENTIAL.
FT DOMAIN 36 225 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 236 310 FZ.
FT DOMAIN 410 677 KRINGLE.
FT NP_BIND 416 424 PROTEIN KINASE.
FT BINDING 442 442 ATP (BY SIMILARITY).
FT ACT_SITE 539 539 ATP (BY SIMILARITY).
FT MOD_RES 565 565 BY SIMILARITY.
FT MOD_RES 569 569 PHOSPHORYLATION (AUTO-) (BY
FT MOD_RES 570 570 SIMILARITY).
FT CARBOHYD 45 45 PHOSPHORYLATION (AUTO-) (BY
FT CARBOHYD 63 63 SIMILARITY).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 685 AA; 526162D27D5FD7C7 CRC64;

Query Match 25.1%; Score 139; DB 1; Length 685;
Best Local Similarity 34.5%; Pred. NO. 6e-08;
Matches 30; Conservative 19; Mismatches 28; Indels 10; Gaps 5;

Oy 1 KTCVEGNGHYRGKASDTMGRCLLPWNSATVLQQTTHAFRSDALQLGLGKKNYCRNPDN 60
Db 235 ENCWEDGSIYRGVANVSASGKCLAW--SWLMKEI-----SDFPEL-IGQ-NYCRNPGS 285

Oy 61 -RRRPWCYOVGLKPLVQECWHDGAD 86
Db 286 VENSPPWCFVDSRRRIELCDIPKCAD 312

RESULT 44
HGF_MOUSE STANDARD; PRT; 728 AA.
ID -HGF_MOUSE STANDARD; PRT; 728 AA.
AC Q08048; Q61662; Q64007;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
DE (Hepatopoietin A).
GN HGF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT), AND SEQUENCE OF 496-504.
RC TISSUE=Mammary fibroblast;
RX MEDLINE=94183257; PubMed=9135822;
RA Sasaki M., Nishio M., Sasaki T., Enami J.;
RT "Identification of mouse mammary fibroblast-derived mammary growth
RT factor as hepatocyte growth factor."
RL Biochem. Biophys. Res. Commun. 199;772-779(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94363381; PubMed=8081873;
RA Lee C.C., Kozak C.A., Yamada K.M.;
RT "Structure, genetic mapping, and expression of the mouse Hgf/scatter
RT factor gene."
RL Cell Adhes. Commun. 1:101-111(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94060105; PubMed=8241272;
RA Liu Y., Michalopoulos G.K., Zarnegar R.;
RT "Molecular cloning and characterization of cDNA encoding mouse
RT hepatocyte growth factor."

```

```

RL CC Biochim. Biophys. Acta 1216:299-303(1993).
CC -!- FUNCTION: HGF is a potent mitogen for mature parenchymal
CC hepatocyte cells, seems to be an hepatotrophic factor, and acts as
CC growth factor for a broad spectrum of tissues and cell types. It
CC has no detectable protease activity.
CC -!- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a
CC disulfide bond.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q08048-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q08048-2; Sequence=VSP_005408;
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -!- SIMILARITY: Contains 4 kringle domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D10212; BAA01064.1; -
CC EMBL; D10213; BAA01065.1; -
CC EMBL; S71816; AAB31855.1; -
CC EMBL; X72307; CAAS1054.1; ALT_INIT.
CC PIR; JC2117; A60185.
CC HSSP; P14210; 1BHT.
CC MESOP; S01_982; -
CC MGD; MGI:96079; Hgf.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR003014; PAN.
CC InterPro; IPR003609; Pan app.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00051; Kringle; 4.
CC Pfam; PF00024; PAN; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 4.
CC SMART; SM00130; KR; 4.
CC SMART; SM00473; PAN AP; 1.
CC SMART; SM00020; Tryp_Sec; 1.
CC PROSITE; PS00021; KRINGLE_1; 4.
CC PROSITE; PS00070; KRINGLE_2; 4.
CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC Growth factor; Kringle; Glycoprotein; Serine protease homolog; Repeat;
CC Signal; Alternative splicing; Pyrrolidone carboxylic acid.
FT SIGNAL 1 32 BY SIMILARITY.
FT CHAIN 33 495 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID
FT MOD_RES 33 33 (BY SIMILARITY).
FT DOMAIN 33 128 PAP.
FT DOMAIN 129 207 KRINGLE 1.
FT DOMAIN 212 289 KRINGLE 2.
FT DOMAIN 306 384 KRINGLE 3.
FT DOMAIN 392 470 KRINGLE 4.
FT DOMAIN 496 728 SERINE PROTEASE-LIKE.
FT DISULFID 71 97 BY SIMILARITY.
FT DISULFID 75 85 BY SIMILARITY.
FT DISULFID 488 607 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 656 656 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPFLC 163 167 Missing (in isoform Short).
FT CONFLICT 344 344 /FTid=VSP_005408.
FT CONFLICT 344 344 N -> K (IN REF. 2).

```

```

FT CONFLICT 479 479 V -> L (IN REF. 2).
FT CONFLICT 564 564 R -> H (IN REF. 3).
SQ SEQUENCE 728 AA; 82944 MW; A0381FC497534328 CRC64;

Query Match 24.8%; Score 137.5; DB 1; Length 728;
Best Local Similarity 32.0%; Pred. No. 9.4e-08;
Matches 31; Conservative 12; Mismatches 35; Indels 19; Gaps 4;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOYTHAHRSDALQLGLG-----KHN YCRNP 58
Db 306 CIQGGEGYRGTSNTIWN GIPQORWDS-----QYPHKHDTIPENFKCKDLRENYCRNP 358

QY 59 DNRERPWCY-----VOVGLKPLVQECMV---HDCADG 87
Db 359 DGAESPWCFTTDPNIRVGVCSPKCDVSSGQDCYRG 395

RESULT 45
HGF RAT
ID - HGF RAT STANDARD; PRT; 728 AA.
AC P17945;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
DE (Hepatopoietin A).
GN HGF
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=90222197; PubMed=2139229;
RA Toshiro K., Hagiya M., Nishizawa T., Seki T., Shimonishi M., Shimizu S., Nakamura T.;
RT "Deduced primary structure of rat hepatocyte growth factor and expression of the mRNA in rat tissues.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3200-3204(1990).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=91031482; PubMed=2146117;
RA Okajima A., Miyazawa K., Kitamura N.;
RT "Primary structure of rat hepatocyte growth factor and induction of its mRNA during liver regeneration following hepatic injury.";
RL Eur. J. Biochem. 193:375-381(1990).
CC -!- FUNCTION: HGF is a potent mitogen for mature parenchymal hepatocyte cells, seems to be an hepatotropic factor, and acts as growth factor for a broad spectrum of tissues and cell types. It has no detectable protease activity.
CC -!- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a disulfide bond.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -!- SIMILARITY: Contains 4 kringle domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC ENBL; D90102; BAA14133.1; -
CC DR ENBL; X54400; CAA38266.1; -
CC DR PIR; A35644; A35644.
CC DR HSSP; P14210; 1BHT.
CC DR MEROPS; S01.978; -.
CC DR InterPro; IPR009003; Cys_Ser_trypsin.
CC DR InterPro; IPR000001; Kringle.
CC DR InterPro; IPR003014; PAN.

```

```

DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Growth factor; Kringle; Glycoprotein; Serine protease homolog; Repeat;
KW signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 32 BY SIMILARITY.
FT CHAIN 33 495 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
FT CHAIN 496 728 HEPATOCYTE GROWTH FACTOR BETA CHAIN.
FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID (BY SIMILARITY).
FT DOMAIN 33 128 PAP.
FT DOMAIN 129 207 KRINGLE 1.
FT DOMAIN 212 289 KRINGLE 2.
FT DOMAIN 306 384 KRINGLE 3.
FT DOMAIN 392 470 KRINGLE 4.
FT DOMAIN 496 728 SERINE PROTEASE-LIKE.
FT DISULFID 71 97 BY SIMILARITY.
FT DISULFID 75 85 BY SIMILARITY.
FT DISULFID 488 607 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 656 656 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 728 AA; 82905 MW; 3E0BF1F96ADCEBFF CRC64;

Query Match 24.5%; Score 135.5; DB 1; Length 728;
Best Local Similarity 32.0%; Pred. No. 1.6e-07;
Matches 31; Conservative 11; Mismatches 36; Indels 19; Gaps 4;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOYTHAHRSDALQLGLG-----KHN YCRNP 58
Db 306 CIQGGEGYRGTSNTIWN GIPQORWDS-----QYPHKHDTIPENFKCKDLRENYCRNP 358

QY 59 DNRERPWCY-----VOVGLKPLVQECMV---HDCADG 87
Db 359 DGAESPWCFTTDPNIRVGVCSPKCDVSSGQDCYRG 395

RESULT 46
HGF HUMAN
ID - HGF HUMAN STANDARD; PRT; 728 AA.
AC P14210; Q9BYL9; Q9UDU6;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
DE (Hepatopoietin A).
GN HGF OR HPFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=91340155; PubMed=1831432;
RA Seki T., Hagiya M., Shimonishi M., Nakamura T., Shimizu S.;
RT "Organization of the human hepatocyte growth factor-encoding gene.";
RL Gene 102:213-219(1991).
RN [2]
SEQUENCE FROM N.A.
RP TISSUE=Placenta;
RC

```

RX MEDLINE=89392017; PubMed=2528952;
RA Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,
RA Arakaki N., Nakayama H., Hirono S., Sakiyama O., Takahashi K.,
RA Gohda E., Daikuhara Y., Kitamura N.,
RT "Molecular cloning and sequence analysis of cDNA for human hepatocyte
growth factor.";
RL Biochem. Biophys. Res. Commun. 163:967-973(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91025062; PubMed=2145836;
RA Seki T., Ihara I., Sugimura A., Shimonishi M., Nishizawa T.,
RA Asami O., Hagiya M., Nakamura T., Shimizu S.,
RT "Isolation and expression of cDNA for different forms of hepatocyte
growth factor from human leukocyte.";
RL Biochem. Biophys. Res. Commun. 172:321-327(1990).
RN [4]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.
RC TISSUE=Liver;
RX MEDLINE=90066676; PubMed=2531289;
RA Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimonishi M.,
RA Sugimura A., Tashiro K., Shimizu S.,
RT "Molecular cloning and expression of human hepatocyte growth factor.";
RL Nature 342:440-443(1989).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryonic fibroblast;
RX MEDLINE=91334393; PubMed=1831266;
RA Weidner K.M., Arakaki N., Hartmann G., Vandekerkhove J., Weingart S.,
RA Rieder H., Fonatsch C., Tsubouchi H., Hishida T., Daikuhara Y.,
RA Birchmeier W.,
RT "Evidence for the identity of human scatter factor and human
hepatocyte growth factor.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).
RN [6]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Courtney L., Elliot G., Angell S.,
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE OF 249-695 FROM N.A.
RX MEDLINE=91369928; PubMed=1832556;
RA Miyazawa K., Kitamura A., Kitamura N.,
RT "Structural organization and the transcription initiation site of the
human hepatocyte growth factor gene.";
RL Biochemistry 30:9170-9176(1991).
RN [9]
RP SIGNAL SEQUENCE CLEAVAGE SITE.
RX MEDLINE=91207365; PubMed=1826837;
RA Yoshiyama Y., Arakaki N., Naka D., Takahashi K., Hirono S., Kondo J.,
RA Nakayama H., Gohda E., Kitamura N., Tsubouchi H., Ishii T.,
RA Hishida T., Daikuhara Y.,
RT "Identification of the N-terminal residue of the heavy chain of both
native and recombinant human hepatocyte growth factor.";
RL Biochem. Biophys. Res. Commun. 175:660-667(1991).
RN [10]
RP CARBOHYDRATE-LINKAGE SITE THR-476.
RX MEDLINE=91129192; PubMed=1482348;
RA Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,
RA Nakamura T., Shimizu S.,
RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide
on the alpha chain.";
RL Biochem. Biophys. Res. Commun. 189:1329-1335(1992).
RN [11]
RP MUTAGENESIS.
RX MEDLINE=92331602; PubMed=1321034;
RA Lokker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,
RA Baker J.B., Godowski P.J.;

"Structure-function analysis of hepatocyte growth factor:
identification of variants that lack mitogenic activity yet retain
high affinity receptor binding.";
RL EMBO J. 11:2503-2510(1992).
RN [12]
RP STRUCTURE BY NMR OF 31-127.
RX MEDLINE=98154323; PubMed=943272;
RA Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,
RA Rubin J.S., Bottaro D.P., Byrd R.A.;
RT "The solution structure of the N-terminal domain of hepatocyte growth
factor reveals a potential heparin-binding site.";
RN Structure 6:109-116(1998).
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.
RX MEDLINE=99036858; PubMed=9817840;
RA Utsch M., Lokker N.A., Godowski P.J., de Vos A.M.;
RT "Crystal structure of the NK1 fragment of human hepatocyte growth
factor at 2.0-A resolution.";
RN Structure 6:1383-1393(1998).
CC CC -!- FUNCTION: HGF is a potent mitogen for mature parenchymal
hepatocyte cells, seems to be an hepatotrophic factor, and acts
as growth factor for a broad spectrum of tissues and cell types.
CC It has no detectable protease activity.
CC -!- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a
disulfide bond.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -!- SIMILARITY: Contains 4 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----

EMBL; D90334; BAA14348.1; -
EMBL; D90318; BAA14348.1; JOINED.
EMBL; D90319; BAA14348.1; JOINED.
EMBL; D90320; BAA14348.1; JOINED.
EMBL; D90322; BAA14348.1; JOINED.
EMBL; D90323; BAA14348.1; JOINED.
EMBL; D90324; BAA14348.1; JOINED.
EMBL; D90325; BAA14348.1; JOINED.
EMBL; D90326; BAA14348.1; JOINED.
EMBL; D90327; BAA14348.1; JOINED.
EMBL; D90328; BAA14348.1; JOINED.
EMBL; D90329; BAA14348.1; JOINED.
EMBL; D90330; BAA14348.1; JOINED.
EMBL; D90331; BAA14348.1; JOINED.
EMBL; D90332; BAA14348.1; JOINED.
EMBL; D90333; BAA14348.1; JOINED.
EMBL; M29145; AAA52650.1; -
EMBL; M60718; AAA52648.1; -
EMBL; X16323; CAA34387.1; -
EMBL; M73239; AAA64239.1; -
EMBL; M73240; AAA64297.1; -
EMBL; AY246560; AAO61091.1; -
EMBL; AC004960; AAC71655.1; -
EMBL; M75983; AAG53460.1; -
EMBL; M75972; AAG53460.1; JOINED.
EMBL; M75973; AAG53460.1; JOINED.
EMBL; M75974; AAG53460.1; JOINED.
EMBL; M75975; AAG53460.1; JOINED.
EMBL; M75976; AAG53460.1; JOINED.
EMBL; M75977; AAG53460.1; JOINED.
EMBL; M75978; AAG53460.1; JOINED.
EMBL; M75979; AAG53460.1; JOINED.
EMBL; M75980; AAG53460.1; JOINED.
EMBL; M75981; AAG53460.1; JOINED.
EMBL; M75982; AAG53460.1; JOINED.
PIR; JH0579; JH0579.
PDB; 2HGf; 24-JUN-98.

```

DR PDB; 1BHT; 18-NOV-98.
DR PDS; INK1; 29-DEC-99.
DR MEROPS; S01.376; -.
DR GlycosultedB; P4210; -.
DR Genew; HGNC:4893; HGF.
DR MIM; 142409; -.
DR GO; GO:0008083; P: growth factor activity; NAS.
DR GO; GO:0007087; P: mitosis; NAS.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRYSP; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS00240; TRYPsin DOM; 1.
DR Growth factor; Kringle; Glycoprotein; Serine protease homolog; Repeat;
KW Signal; 3D-structure; Polymorphism; Pyrrolidone carboxylic acid.
FT SIGNAL 1 31
FT CHAIN 32 494
FT CHAIN 495 728
FT MOD RES 32 32
FT DOMAIN 32 127
FT PYRROLIDONE CARBOXYLIC ACID.

Query Match 24.1%; Score 133.5; DB 1; Length 728;
Best Local Similarity 34.7%; Pred No. 2.7e-07;
Matches 33; Conservative 10; Mismatches 37; Indels 15; Gaps 5;

Qy 3 CYEGNGHFRGKASTDTMGRLPCLPWSATVLOQTYHAH--RSDALQLGLGKHNYCRNDPN 60
Db 305 CIGQGGEGYRGVTWINGIPQRWDS----QYPHEHMTPTENFKCDLRENYCRNDPG 359
Qy 61 RRRPWCY-----VQGLKPLVQEC-MVH--DCADG 87
Db 360 SESFPCFTTDPNIRVGYCSQIENCDMSHQDQCYRG 394

RESULT 47
ID PLMN PETWA STANDARD; PRT; 325 AA.
AC F33574;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragments).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hypercoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OC NCBI_TaxID=7757;
RN [1]
RP SEQUENCES.
RA Affolter M., Schaller J., Rickli E.E.;
RT "Isolation, characterization and partial amino acid sequence of
RT lamprey plasminogen."
RL Protein Seq. Data Anal. 5:207-211(1993).
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
CC a proteolytic factor in a variety of other processes including
CC embryonic development, tissue remodeling, tumor invasion,
CC and inflammation.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.

```

```

CC -!- SIMILARITY: Contains at least 2 kringle domains.
DR PIR; S33879; S33879.
DR HSSP; P00747; SHPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; TRYSP; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPsin DOM; PARTIAL.
DR PROSITE; PS00134; TRYPsin HIS; PARTIAL.
DR PROSITE; PS00135; TRYPsin_SER; PARTIAL.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT NON_CONS 15 16
FT NON_CONS 34 35
FT NON_CONS 44 45
FT NON_CONS 76 77
FT NON_CONS 111 112
FT NON_CONS 138 139
FT NON_CONS 158 159
FT NON_CONS 178 179
FT NON_CONS 216 217
FT NON_CONS 236 237
FT NON_CONS 267 268
FT NON_CONS 282 283
FT NON_CONS 295 296
FT NON_CONS 307 308
FT NON_CONS 315 316
FT NON_TER 325 325
SQ SEQUENCE 325 AA; 1550B539AC6ED3C CRC64;

Query Match 23.6%; Score 131; DB 1; Length 325;
Best Local Similarity 32.9%; Pred. No. 2.2e-07;
Matches 28; Conservative 5; Mismatches 36; Indels 16; Gaps 2;

Qy 3 CYEGNGHFRGKASTDTMGRLPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNDNR 62
Db 81 CVKGTGEGYRGTAALTVSGKACQAWASQT-----PDGVYSCQGLVSNYCRNPDGK 131
Qy 63 RPMCYVQVGLKPLVQECMVHDCADG 87
Db 132 LPWCYT-----TEYCNVPSCTGG 149

RESULT 48
ID THRB MOUSE STANDARD; PRT; 618 AA.
AC P12221;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Prothrombin precursor (EC 3.4.21.5).
GN F2 OR CF2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Liver;
RX MEDLINE=91025553; PubMed=2222810;
RA Fritzer Degen S.J., Schaffer L.A., Jamison C.S., Grant S.G.,
RA Fitzgibbon J.J., Pai J.-A., Chapman V.M., Elliott R.W.;
RT "Characterization of the cDNA coding for mouse prothrombin and
RT localization of the gene on mouse chromosome 2.";
RL DNA Cell Biol. 9:487-498(1990).
RN [2]
RP SEQUENCE FROM N.A.

```


RC STRAIN=FVB/N; TISSUE=Liver;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.L., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrincci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.C.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP SEQUENCE OF 384-618 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=92212913; PubMed=1557383;
 RA Banfield D.K., Macgillivray R.T.;
 RA "Partial characterization of vertebrate prothrombin cDNAs:
 RT amplification and sequence analysis of the B chain of thrombin from
 RT nine different species";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783 (1992).
 CC -|- FUNCTION: Thrombin, which cleaves bonds after Arg and Lys,
 CC converts fibrinogen to fibrin and activates factors V, VII, VIII,
 CC XI, and, in complex with thrombomodulin, protein C.
 CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
 CC fibrinogen to fibrin and releases fibrinopeptide A and B.
 CC -|- PFM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
 CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOBIAL
 CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
 CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
 CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION
 CC OF PROTHROMBIN TO THROMBIN.
 CC -|- MISCELLANEOUS: Prothrombin is activated on the surface of a
 CC phospholipid membrane that binds the amino end of prothrombin and
 CC factors Va and Xa in Ca-dependent interactions; factor Xa removes
 CC the activation peptide and cleaves the remaining part into light
 CC and heavy chains. The activation process starts slowly because
 CC factor V itself has to be activated by the initial, small amounts
 CC of thrombin.
 CC -|- MISCELLANEOUS: Thrombin can itself cleave the amino terminal
 CC fragment (fragment 1) of the prothrombin, prior to its activation
 CC by factor Xa.
 CC -|- SIMILARITY: Belongs to peptidase family S1.
 CC -|- SIMILARITY: Contains 2 kringle domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X52308; CAA36548.1; -;
 CC EMBL; BC013662; AAH13662.1; -;
 CC EMBL; M81394; AAA40435.1; -;
 CC F01; A35827; A35827.
 CC HSPF; P00734; 1B7X.
 CC MEROPS; S01.217; -;
 CC MGD; MGI:88380; F2.
 CC InterPro; IPR009003; Cys Ser trypsin.
 CC InterPro; IPR002383; GLA_blood.
 CC InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR003966; Peptidase_S1A_pr.
 DR InterPro; IPR000294; VitK_dep_GLA.
 DR Pfam; PF00594; gla; 1.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00001; GLABLOOD.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR01505; PROTHROMBIN.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00069; GLA; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00200; TRYPSIN; 1.
 DR PROSITE; PS00011; GLU CARBOXYLATION; 1.
 DR PROSITE; PS00021; KRINGLE 1; 2.
 DR PROSITE; PS00070; KRINGLE 2; 2.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
 KW Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
 KW Hydrolase; Serine protease; Kringle; Signal.
 FT SIGNAL 1 24
 FT PROPEP 25 43
 FT CHAIN 44 618
 FT PEPTIDE 44 200
 FT PEPTIDE 201 324
 FT CHAIN 325 360
 FT CHAIN 361 618
 FT DOMAIN 109 187
 FT DOMAIN 215 292
 FT DOMAIN 361 618
 FT SITE 200 201
 FT SITE 324 325
 FT SITE 360 361
 FT ACT_SITE 403 403
 FT ACT_SITE 459 459
 FT ACT_SITE 565 565
 FT MOD_RES 50 50
 FT MOD_RES 51 51
 FT MOD_RES 58 58
 FT MOD_RES 60 60
 FT MOD_RES 63 63
 FT MOD_RES 64 64
 FT MOD_RES 69 69
 FT MOD_RES 70 70
 FT MOD_RES 73 73
 FT MOD_RES 76 76
 FT DISULFID 91 104
 FT DISULFID 109 187
 FT DISULFID 130 170
 FT DISULFID 158 182
 FT DISULFID 215 293
 FT DISULFID 236 276
 FT DISULFID 264 288
 FT DISULFID 333 479
 FT DISULFID 388 404
 FT DISULFID 533 547
 FT DISULFID 551 591
 FT CARBOHYD 122 122
 FT CARBOHYD 144 144
 FT CARBOHYD 413 413
 FT CARBOHYD 553 553
 SQ SEQUENCE 618 AA; 70268 MW; B89F719A9AFD601E0 CRC64;

Query Match 22.7%; Score 125.5; DB 1; Length 618;
 Best Local Similarity 31.8%; Pred. No. 1.8e-06;
 Matches 27; Conservative 13; Mismatches 40; Indels 5; Gaps 3;

QY 3 CYEGNGHYRGKASDTWGRCLPNSATVLQQTVAHRSALQLGLGKHVCNRPD-NR 61

```
Db 215 CUTEGRLYQGNLAVTTIGSPCLPWNLSIPAKTSLKYQDFPEVKL---VENFCRNPDWDE 271
QY 62 RRPWCYVQVGLKPLVQECNMVHDCAD 86
Db 272 EGACVCV-AGQGDPEYCNLYCEE 295

RESULT 49
PLMN_SHEEP
ID PLMN_SHEEP STANDARD; PRT; 343 AA.
AC P81286;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragment).
GN PLG.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RX MEDLINE=93149995; PubMed=1492092;
RA Schaller J., Straub C., Kamper U., Rickli E.E.;
RT "Complete amino acid sequence of ovine miniplasminogen.";
RL Protein Seq. Data Anal. 5:21-25(1992).
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
CC a proteolytic factor in a variety of other processes including
CC embryonic development, tissue remodeling, tumor invasion, and
CC inflammation; in ovulation it weakens the walls of the Graafian
CC follicle. It activates the urokinase-type plasminogen activator,
CC collagenases and several complement zymogens, such as C1 and C5.
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von
CC Willebrand factor.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen
CC activators, both plasminogen and its activator being bound to
CC fibrin. Cannot be activated with streptokinase.
CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
CC immediately after dissociation from the clot plasminogen subfamily.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -!- SIMILARITY: Contains at least 2 kringle domains.
DR PIR; B61545; B61545.
DR HSSP; P00747; 5HPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.
FT NON TER 1 1
FT DOMAIN <1 140 HEAVY CHAIN A.
FT DOMAIN 141 >343 LIGHT CHAIN A.
FT DOMAIN <1 17 KRINGLE 4.
FT DOMAIN
```

```
FT DOMAIN 41 120 KRINGLE 5.
FT DOMAIN 114 341 SERINE PROTEASE.
FT ACT_SITE 181 CHARGE RELAY SYSTEM.
FT ACT_SITE 224 CHARGE RELAY SYSTEM.
FT ACT_SITE 319 CHARGE RELAY SYSTEM.
FT NON_TER 343
SQ SEQUENCE 343 AA; 8DF6EBA92D596EE0 CRC64;

Query Match 22.5%; Score 124.5; DB 1; Length 343;
Best local similarity 35.3%; Pred. No. 1.2e-06;
Matches 36; Conservative 8; Mismatches 43; Indels 15; Gaps 6;

QY 3 CYEGNGHFVRGKASTDTMGRCPLPWNLSATVLOQTYHAH---RSDALQLGLGHNYCRNP 58
Db 15 CMLGIGKYGKATTVAGVPCQEWAA-----QEPHKGIFPTETNPRAGLEK-NYCRNP 68
QY 59 D-NRRRPWCYVQVGLKPLVQECNMVHDCAD---DGKPPSSPPEE 96
Db 69 DGDVNGVWCYT-TNPKLFYCDIPQCESSFDGKPKVEPKK 109

RESULT 50
THRB_HUMAN
ID THRB_HUMAN STANDARD; PRT; 622 AA.
AC P00734;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Prothrombin precursor (EC 3.4.21.5) (Coagulation factor II).
GN F2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88077877; PubMed=2825773;
RA Degen S.J.F., Davie E.W.;
RT "Nucleotide sequence of the gene for human prothrombin.";
RL Biochemistry 26:6165-6177(1987).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT MET-165.
RA Rieder M.J., Arnel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA Osuna M., Peel C.L., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 8-622 FROM N.A.
RX MEDLINE=83231469; PubMed=6305407;
RA Degen S.J.F., McGallivray R.T.A., Davie E.W.;
RT "Characterization of the complementary deoxyribonucleic acid and gene
RL coding for human prothrombin.";
RL Biochemistry 22:2087-2097(1983).
RN [4]
RP SEQUENCE OF 44-314.
RX MEDLINE=77193964; PubMed=266717;
RA Walz D.A., Hewett-Emmett D., Sesegs W.H.;
RT "Amino acid sequence of human prothrombin fragments 1 and 2.";
RL Proc. Natl. Acad. Sci. U.S.A. 74:1969-1972(1977).
RN [5]
RP SEQUENCE OF 315-622.
RX MEDLINE=7207112; PubMed=873923;
RA Butkowski R.J., Eliot J., Downing M.R., Mann K.G.;
RT "Primary structure of human prothrombin 2 and alpha-thrombin.";
RL J. Biol. Chem. 252:4942-4957(1977).
RN [6]
RP PROCESSING.
RX MEDLINE=87008532; PubMed=3759958;
RA Rabiet M.J., Blashill A., Furie B., Furie B.C.;
RT "Prothrombin fragment 1 X 2 X 3, a major product of prothrombin
RL activation in human plasma.";
RL J. Biol. Chem. 261:13210-13215(1986).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
```

RP MEDLINE=90059942; PubMed=2503108;
RA Bode W., Mayr I., Baumann U., Huber R., Stone S.R., Hofsteenge J.;
RT "The refined 1.9 A crystal structure of human alpha-thrombin;
RT interaction with D-Phe-Pro-Arg chloromethylketone and significance of
RT the Tyr-Pro-Pro-Tyr insertion segment.";
RL EMBO J. 8:3467-3475 (1989).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RP MEDLINE=90327074; PubMed=2374926;
RA Rydel T.J., Ravichandran K.G., Tulinsky A., Bode W., Huber R.,
RA Roitsch C., Fenton J.W. II;
RT "The structure of a complex of recombinant hirudin and human alpha-
RT thrombin.";
RL Science 249:277-280 (1990).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RP MEDLINE=94350942; PubMed=8071320;
RA Rydel T.J., Yin M., Padmanabhan K.P., Blankenship D.T., Cardin A.D.,
RA Correa P.E., Fenton J.W. II, Tulinsky A.;
RT "Crystallographic structure of human gamma-thrombin.";
RL J. Biol. Chem. 269:22000-22006 (1994).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RP MEDLINE=97357286; PubMed=9214615;
RA van de Loch A., Bode W., Huber R., le Bonniec B.F., Stone S.R.,
RA Esmen C.T., Stubbs M.T.;
RT "The thrombin E192Q-BPTI complex reveals gross structural
RT rearrangements: implications for the interaction with antithrombin
RT and thrombomodulin.";
RL EMBO J. 16:2977-2984 (1997).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 328-601.
RP MEDLINE=99162521; PubMed=10051558;
RA Guinto E.R., Caccia S., Rose T., Fuetterer K., Wakeman G., di Cera E.;
RT "Unexpected crucial role of residue 225 in serine proteases.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:1852-1857 (1999).
RN [12]
RP VARIANT DYSPROTHROMBINEMIA CYS-314.
RP MEDLINE=87033739; PubMed=3771562;
RA Rabiet M.-J., Furie B.C., Furie B.;
RT "Molecular defect of prothrombin Barcelona. Substitution of cysteine
RT for arginine at residue 273.";
RL J. Biol. Chem. 261:15045-15048 (1986).
RN [13]
RP VARIANT DYSPROTHROMBINEMIA ALA-509.
RP MEDLINE=95313001; PubMed=7792730;
RA Degen S.J.F., McDowell S.A., Sparks L.M., Scharer I.;
RT "Prothrombin Frankfurt: a dysfunctional prothrombin characterized by
RT substitution of Glu-466 by Ala.";
RL Thromb. Haemost. 73:203-209 (1995).
RN [14]
RP VARIANTS DYSPROTHROMBINEMIA THR-380 AND HIS-431.
RP MEDLINE=93043342; PubMed=1421398;
RA Morishita E., Saito M., Kumabashiri I., Asakura H., Matsuda T.,
RA Yamaguchi K.;
RT "Prothrombin Hmi: a compound heterozygote for two dysfunctional
RT prothrombin molecules (Met-337-->Thr and Arg-388-->His).";
RL Blood 80:2275-2280 (1992).
RN [15]
RP VARIANT DYSPROTHROMBINEMIA HIS-314.
RP MEDLINE=95169898; PubMed=7865694;
RA James H.L., Kim D.J., Zheng D.-Q., Girolami A.;
RT "Prothrombin Padua I: incomplete activation due to an amino acid
RT substitution at a factor Xa cleavage site.";
RL Blood Coagul. Fibrinolysis 5:841-844 (1994).
RN [16]
RP VARIANT DYSPROTHROMBINEMIA CYS-425.
RP MEDLINE=89207504; PubMed=3242619;
RA Henriksen R.A., Mann K.G.;
RT "Identification of the primary structural defect in the dysfibrinogen
RT thrombin Quick I: substitution of cysteine for arginine-382.";
RL Biochemistry 27:9160-9165 (1988).
RN [17]
RP VARIANT DYSPROTHROMBINEMIA VAL-601.
RP MEDLINE=89247398; PubMed=2719946;
RA Henriksen R.A., Mann K.G.;
RT "Substitution of valine for glycine-558 in the congenital dysfibrinogen
RT thrombin Quick II alters primary substrate specificity.";
RL Biochemistry 28:2078-2082 (1989).
RN [18]
RP VARIANT DYSPROTHROMBINEMIA ALA-509.
RP MEDLINE=92378975; PubMed=1354985;
RA Miyata T., Aruga R., Uneyama H., Bezeaud A., Guillin M.-C.,
RA Iwanaga S.;
RT "Prothrombin Salakta: substitution of glutamic acid-466 by alanine
RT reduces the fibrinogen clotting activity and the esterase activity.";
RL Biochemistry 31:7457-7462 (1992).
RN [19]
RP VARIANT DYSPROTHROMBINEMIA TRP-461.
RP MEDLINE=87185407; PubMed=3567158;
RA Miyata T., Morita T., Inomoto T., Kawauchi S., Shirakami A.,
RA Iwanaga S.;
RT "Prothrombin Tokushima, a replacement of arginine-418 by tryptophan
RT that impairs the fibrinogen clotting activity of derived thrombin
RT Tokushima.";
RL Biochemistry 26:1117-1122 (1987).
RN [20]
RP VARIANT DYSPROTHROMBINEMIA TRP-461.
RP MEDLINE=87101511; PubMed=3801671;
RA Inomoto T., Shirakami A., Kawauchi S., Shigeakiyo T., Saito S.,
RA Miyoshi K., Morita T., Iwanaga S.;
RT "Prothrombin Tokushima: characterization of dysfunctional thrombin
RT derived from a variant of human prothrombin.";
RL Blood 69:565-569 (1987).
RN [21]
RP VARIANT DYSPROTHROMBINEMIA TRP-461.
RP MEDLINE=92256895; PubMed=1349838;
RA Iwahana H., Yoshimoto K., Shigeakiyo T., Shirakami A., Saito S.,
RA Itakura M.;
RT "Detection of a single base substitution of the gene for prothrombin
RT Tokushima. The application of PCR-SSCP for the genetic and molecular
RT analysis of dysprothrombinemia.";
RL Int. J. Hematol. 55:93-100 (1992).
RN [22]
RP VARIANT DYSPROTHROMBINEMIA LYS-200.
RP MEDLINE=83204687; PubMed=6405779;
RA Board P.G., Shaw D.C.;
RT "Determination of the amino acid substitution in human prothrombin
RT type 3 (157 Glu leads to Lys) and the localization of a third
RT thrombin cleavage site.";
RL Br. J. Haematol. 54:245-254 (1983).
RN [23]
RP VARIANTS MET-165 AND THR-386.
RP MEDLINE=99318093; PubMed=10391209;
RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemes J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
RT of human genes.";
RL Nat. Genet. 22:231-238 (1999).
RN [24]
RP ERRATUM.
RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemes J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
RA Lander E.S.;
RT "FUNCTION: Thrombin, which cleaves bonds after Arg and Lys,
RT converts fibrinogen to fibrin and activates factors V, VII, VIII,
RT XIII, and, in complex with thrombomodulin, protein C.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
CC fibrinogen to fibrin and releases fibrinopeptide A and B.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
CC -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,

CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC ENZYME. THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION

Query Match 22.4%; Score 124; DB 1; Length 622;
Best Local Similarity 38.8%; Pred. No. 2.6e-06;
Matches 26; Conservative 8; Mismatches 29; Indels 4; Gaps 2;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQQTVAHRS DALQLGLGKHNYCRNPD-NR 61
Db 213 CVPDRGQYQGR LAVTTHTGLPCLAWASQAQKALS KHD FNSAVQL---VENFCRNPDGDE 269
QY 62 RRPWCYV 68
Db 270 EGVWCYV 276

Search completed: May 25, 2004, 14:53:57
Job time : 4.70202 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:44:05 ; Search time 15.9018 Seconds
(without alignments)
1904.795 Million cell updates/sec

Title: US-09-880-503-9
Perfect score: 554
Sequence: 1 KTCVGNHFRGKASTDTM.....QECMVHDCADGKFPSPPEE 96

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

- Database :
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriapi.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	495	89.4	154	4 Q96SE8	Q96SE8 homo sapien
2	440	79.4	433	6 Q8MIL0	Q8MIL0 oryctolagus
3	440	79.4	433	6 Q8MHV7	Q8MHV7 oryctolagus
4	422	76.2	157	6 Q9TVB8	Q9TVB8 bos taurus
5	408	73.6	231	11 Q8C6L2	Q8C6L2 mus musculus
6	337.5	60.9	214	6 Q9XIT70	Q9XIT70 oryctolagus
7	236.5	42.7	564	6 Q8MKB1	Q8MKB1 oryctolagus
8	226	40.8	291	4 Q7Z7N2	Q7Z7N2 homo sapien
9	226	40.8	516	4 Q9BU99	Q9BU99 homo sapien
10	226	40.8	562	4 Q8GVK8	Q8GVK8 homo sapien
11	218.5	39.4	90	4 Q8NG20	Q8NG20 homo sapien
12	218.5	39.4	395	4 Q9BZW1	Q9BZW1 homo sapien
13	211	38.1	562	6 Q8SQ23	Q8SQ23 sus scrofa
14	210	37.9	517	11 Q8KOD2	Q8KOD2 mus musculus
15	205	37.0	540	13 Q800Y7	Q800Y7 meleagris g
16	199.5	36.0	560	4 Q14520	Q14520 homo sapien

17	191	34.5	202	13	Q90675	Q90675 gallus gall
18	188	33.9	653	11	Q8VCS4	Q8VCS4 mus musculus
19	187.5	33.8	597	11	Q35727	Q35727 mus musculus
20	187.5	33.8	609	11	Q80YC5	Q80YC5 mus musculus
21	178.5	32.2	616	6	Q97507	Q97507 sus scrofa
22	172.5	31.1	421	13	Q8AXX3	Q8AXX3 xenopus lae
23	170.5	30.8	615	4	Q81ZZ5	Q81ZZ5 homo sapien
24	167	30.1	947	13	Q8AXY6	Q8AXY6 gallus gall
25	166	30.0	263	13	Q7SXB3	Q7SXB3 brachydanio
26	158	28.5	454	6	Q46506	Q46506 papio hamad
27	157	28.3	113	4	Q9UIR5	Q9UIR5 homo sapien
28	154	27.8	806	6	Q18783	Q18783 macropus eu
29	153.5	27.7	801	11	Q8K0Q8	Q8K0Q8 mus musculus
30	153.5	27.7	944	11	Q8C3W2	Q8C3W2 mus musculus
31	153.5	27.7	944	11	Q8BSP6	Q8BSP6 mus musculus
32	153	27.6	113	4	Q9UIR7	Q9UIR7 homo sapien
33	151.5	27.3	359	6	Q8WNR1	Q8WNR1 canis famli
34	150	27.1	812	11	Q9R0W3	Q9R0W3 rattus norv
35	149	26.9	716	13	Q91691	Q91691 xenopus lae
36	148	26.7	105	4	Q9UIR8	Q9UIR8 homo sapien
37	147.5	26.6	567	4	Q13208	Q13208 homo sapien
38	147.5	26.6	648	4	Q9HIV4	Q9HIV4 homo sapien
39	147.5	26.6	716	11	Q9IXG8	Q9IXG8 mus musculus
40	147	26.5	113	4	Q9UIR6	Q9UIR6 homo sapien
41	146.5	26.4	810	4	Q15146	Q15146 homo sapien
42	146	26.4	393	4	Q9BRB6	Q9BRB6 mus musculus
43	146	26.4	937	11	Q8BNP9	Q8BNP9 mus musculus
44	146	26.4	937	11	Q8BG10	Q8BG10 mus musculus
45	145.5	26.3	313	13	Q9PU78	Q9PU78 crocodylus
46	145	26.2	902	5	Q17576	Q17576 caenorhabdi
47	145	26.2	928	5	Q9BLY1	Q9BLY1 caenorhabdi
48	143.5	25.9	709	13	Q90ZN6	Q90ZN6 brachydanio
49	143	25.8	2869	6	Q28398	Q28398 erinaceus e
50	142.5	25.7	213	11	Q81123	Q81123 mus musculus
51	142.5	25.7	264	11	Q81122	Q81122 mus musculus
52	142.5	25.7	264	11	Q7TMJ8	Q7TMJ8 mus musculus
53	141	25.5	759	11	Q7TP84	Q7TP84 rattus norv
54	140	25.3	132	4	Q16609	Q16609 homo sapien
55	140	25.3	930	13	Q8AV69	Q8AV69 xenopus lae
56	138.5	25.0	704	13	Q90865	Q90865 gallus gall
57	138.5	25.0	709	13	Q7ZTN9	Q7ZTN9 xenopus lae
58	138.5	25.0	717	13	P70006	P70006 xenopus lae
59	138	24.9	1145	5	Q9BKL8	Q9BKL8 aplysia cal
60	137.5	24.8	716	11	P70521	P70521 rattus norv
61	137.5	24.8	728	11	Q8C9G5	Q8C9G5 mus musculus
62	137	24.7	334	6	Q46507	Q46507 papio hamad
63	135.5	24.5	553	13	Q7Z292	Q7Z292 brachydanio
64	135	24.4	429	13	Q8AVB0	Q8AVB0 brachydanio
65	134.5	24.3	730	6	Q867B7	Q867B7 canis famli
66	133.5	24.1	145	6	Q28911	Q28911 macaca fasc
67	133.5	24.1	215	13	Q42341	Q42341 gallus gall
68	133.5	24.1	726	13	Q90978	Q90978 gallus gall
69	131.5	23.7	234	4	Q86YW2	Q86YW2 homo sapien
70	131.5	23.7	263	4	Q00318	Q00318 homo sapien
71	131.5	23.7	263	4	Q8NCU9	Q8NCU9 homo sapien
72	131.5	23.7	263	4	Q96PE7	Q96PE7 homo sapien
73	131	23.6	95	4	Q8N656	Q8N656 homo sapien
74	129	23.3	594	5	P91823	P91823 caenorhabdi
75	127.5	23.0	710	13	Q91402	Q91402 xenopus. he

ALIGNMENTS

RESULT 1
Q96SE8 PRELIMINARY; PRT: 154 AA.
ID Q96SE8
AC Q96SE8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Urokinase-type plasminogen activator amino-terminal fragment.
GN ATF.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu J., Bai X., Ruan C.;
RT "Cloning and expression of the amino-terminal fragment of human
RT urokinase-type plasminogen activator.", to the EMBL/GenBank/DBJ databases.
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Overexpression of the amino-terminal fragment of human urokinase-type
RT plasminogen activator in breast cancer cells results in decreased
RT tumor invasion, growth and angiogenesis.",
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029537; AAK38734.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PIRSF; PIRSF001144; Urk_plasm_act; 1.
KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CF2FCFF505572 CRC64;

Query Match 89.4%; Score 495; DB 4; Length 154;
Best Local Similarity 98.9%; Pred. No. 1.7e-50;
Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFRGKASTDTMGRCPLPWN SATVLOQTYHAHRS DALQLGLGKHN YCRNP DN 60
DB 68 KTCYEGNGHFRGKASTDTMGRCPLPWN SATVLOQTYHAHRS DALQLGLGKHN YCRNP DN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADG 154

RESULT 2
QSMILO
ID QSMILO PRELIMINARY; PRT; 433 AA.
AC QSMILO;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator.
GN PLAU.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22155945; PubMed=12149463;
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA Dichek D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
RT development causes arterial constriction and lumen loss, and
RT accelerates lesion growth.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670 (2002).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY122285; AAM83187.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.

DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004235; F:trypsin activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR008293; Pept_S1A_UPA.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PIRSF; PIRSF001144; Urk_plasm_act; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65E64F36415549B0 CRC64;

Query Match 79.4%; Score 440; DB 6; Length 433;
Best Local Similarity 78.7%; Pred. No. 1.7e-43;
Matches 74; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFRGKASTDTMGRCPLPWN SATVLOQTYHAHRS DALQLGLGKHN YCRNP DN 60
DB 70 KTCYEGNGHFRGKASTDTMGRCPLPWN SATVLOQTYHAHRS DALQLGLGKHN YCRNP DN 129

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPP 94
DB 130 QRRPWCYVQVGLKPLVQECMVHDCADGSSGKKPALPP 163

RESULT 3
QSMHY7
ID QSMHY7 PRELIMINARY; PRT; 433 AA.
AC QSMHY7;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.,
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT complete cds.",
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029517; AAK40239.1; -.
DR EMBL; AB087224; BAC02685.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.

```

DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR01254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR008293; Pept S1A_UPA.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PIRSF; PIRSF001144; Urk_plasm.act; 1.
DR Glycoprotein; Hydrolase_Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6D35A371010A6EE CRC64;

Query Match 79.4%; Score 440; DB 6; Length 433;
Best Local Similarity 78.7%; Pred. No. 1.7e-43;
Matches 74; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
70 KTCYHGDSYRGKANTDMDRPLCLAWNSANVLTXYAHRPDALQLGLGKHNYCRNPDH 129

QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKPKSPPE 94
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130 QRRPCYVQVGLKPLVQECMVHDCADGKPKSPPE 163

RESULT 4
Q9TVAS PRELIMINARY; PRT; 157 AA.
ID Q9TVAS;
AC Q9TVAS;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP TISSUE=Skeletal muscle;
RC MEDLINE=21071388; PubMed=11204721;
RX Balcerzak D., Querengesser L., Dixon W.T., Baracos V.E.;
RA "Coordinate expression of matrix-degrading proteinases and their
RT activators and inhibitors in bovine skeletal muscle.";
RL J. Anim. Sci. 79:94-107(2001).
CC 1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
ENBL; AF144761; AAD30301.1; -.
DR HSP; P00749; IURK.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR Glycoprotein; Kinase; Kringle.
KW NON_TER 1
FT NON_TER 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

InterPro; IPR006209; EGF like.
InterPro; IPR000001; Kringle.
InterPro; IPR01254; Peptidase S1.
InterPro; IPR001314; Peptidase S1A.
InterPro; IPR008293; Pept S1A_UPA.
Pfam; PF00051; Kringle; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; KRINGLE.
ProDom; PD000395; Kringle; 1.
SMART; SM00130; KR; 1.
SMART; SM00020; Tryp SPC; 1.
PROSITE; PS00022; EGF_1; 1.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS00070; KRINGLE_2; 1.
PROSITE; PS00240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
PIRSF; PIRSF001144; Urk_plasm.act; 1.
Glycoprotein; Hydrolase_Kinase; Kringle; Protease; Serine protease.
SEQUENCE 433 AA; 48444 MW; 6D35A371010A6EE CRC64;

Query Match 79.4%; Score 440; DB 6; Length 433;
Best Local Similarity 78.7%; Pred. No. 1.7e-43;
Matches 74; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
70 KTCYHGDSYRGKANTDMDRPLCLAWNSANVLTXYAHRPDALQLGLGKHNYCRNPDH 129

QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKPKSPPE 94
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130 QRRPCYVQVGLKPLVQECMVHDCADGKPKSPPE 163

RESULT 4
Q9TVAS PRELIMINARY; PRT; 157 AA.
ID Q9TVAS;
AC Q9TVAS;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP TISSUE=Skeletal muscle;
RC MEDLINE=21071388; PubMed=11204721;
RX Balcerzak D., Querengesser L., Dixon W.T., Baracos V.E.;
RA "Coordinate expression of matrix-degrading proteinases and their
RT activators and inhibitors in bovine skeletal muscle.";
RL J. Anim. Sci. 79:94-107(2001).
CC 1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
ENBL; AF144761; AAD30301.1; -.
DR HSP; P00749; IURK.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR Glycoprotein; Kinase; Kringle.
KW NON_TER 1
FT NON_TER 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match 76.2%; Score 422; DB 6; Length 157;
Best Local Similarity 75.0%; Pred. No. 7.4e-42;
Matches 72; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
36 KTCYQNGHSYRGKANRDLGRPLCLAWDSFTVLLKMYHAHRSDALQLGLGKHNYCRNPDN 95

QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKPKSPPE 96
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 QRRPCYVQVGLKPLVQECMVHDCADGKPKSPPREK 131

RESULT 5
Q8C6L2 PRELIMINARY; PRT; 231 AA.
ID Q8C6L2;
AC Q8C6L2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR ENBL; AK054349; BAC35743.1; -.
DR PIR; P10534; P10534.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR NON_TER 231
FT NON_TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25E980A682737F2 CRC64;

Query Match 73.6%; Score 408; DB 11; Length 231;
Best Local Similarity 70.8%; Pred. No. 5.1e-40;
Matches 68; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 KTCYHGNGDSYRGKANTDTKGRPLCLAWNAPVLQKPYNAHRPDAISLGLGKHNYCRNPDN 128

QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKPKSPPE 96
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
129 QRRPCYVQVGLKPLVQECMVHDCADGKPKSPSSVDQ 164

RESULT 6
Q9XT70 PRELIMINARY; PRT; 214 AA.
ID Q9XT70;
AC Q9XT70;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Oryctolagus cuniculus (Rabbit)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Yin J., Idell S.;
RT "Partial mRNA of rabbit uPA.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097647; AAD39351.1; -.
DR HSP; P00749; 1BJN.
DR MEROPS; S01_231; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00039; Kringle; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
FT NON TER 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match 60.9%; Score 337.5; DB 6; Length 214;
Best Local Similarity 78.7%; Pred. No. 9.8e-32;
Matches 59; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

Qy 20 MGRCLPWNATVLPQOYTHAHRSDALQGLGKHNCRPNRRPWCYQVGLKPLVQEC 79
Db 1 MDRCLAWNANVLTKYTHAHRSDALQGLGKHNCRPNRPHQRPPWCYQVGLKPLVQEC 60

Qy 80 MVHDCADGKPPSP 94
Db 61 KVHD-SSGKPPALPP 74

RESULT 7
Q8MKBL ID Q8MKBL PRELIMINARY; PRT; 564 AA.
AC Q8MKBL
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tissue-type plasminogen activator.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AY029518; AAK40240.1; -.
DR GO; GO:0005576; C:extracellular; IEA.

```

```

DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; Kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00131; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
KW SEQUENCE 564 AA; 62726 MW; 459D8BAC6D4A937C CRC64;

Query Match 42.7%; Score 236.5; DB 6; Length 564;
Best Local Similarity 46.8%; Pred. No. 2.4e-19;
Matches 44; Conservative 12; Mismatches 37; Indels 1; Gaps 1;

Qy 3 CYEGNGHPYRGKASTDTMGREFCLPWNATVLPQOYTHAHRSDALQGLGKHNCRPNRR 62
Db 217 CYLGKQAYRGTHTLTSGASCLPWNATVLPQOYTHAHRSDALQGLGKHNCRPNRR 62

Qy 63 RPWCYQVGLKPLVQECMVHDCAD-GKKPPSPPE 95
Db 277 KPWCYVGLKPLVQECMVHDCAD-GKKPPSPPE 310

RESULT 8
Q7Z7N2 ID Q7Z7N2 PRELIMINARY; PRT; 291 AA.
AC Q7Z7N2
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen activator, tissue type isoform 2.
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., da Ponte S.H., Hastings N.C.,
RA Ahearn M.O., Kuldane S.A., Rajkumar N., Toth E.J., Yi Q.,
RA Nickerson D.A.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY291060; AAP34246.1; -.
DR SEQUENCE 291 AA; 32191 MW; 874E38C52F50EFID CRC64;

Query Match 40.8%; Score 226; DB 4; Length 291;
Best Local Similarity 47.7%; Pred. No. 2e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

```


QY 2 TCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPDR 61
 Db 126 TCYEDQGISYRGTSWSTAESGAECTWNSSALAKPYSGRRPDALRLGLGNHNYCRNPDRD 185
 QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
 Db 186 SKPCYVFKAGYSSEFCSTPACSEG 211

RESULT 9
 Q9BU99 PRELIMINARY; PRT; 516 AA.
 ID Q9BU99
 AC Q9BU99;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to plasminogen activator, tissue.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Skin;
 RA Strausberg R.;
 RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DR EMBL; BC002795; AA02795.1; -.
 DR HSSP; P00750; IASH.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00008; EGF_1.
 DR Pfam; PF00051; Kringle; 2.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS00070; TRYPSIN_DOM; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;

Query Match 40.8%; Score 226; DB 4; Length 516;
 Best Local Similarity 47.7%; Pred. No. 3.8e-18;
 Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPDR 61
 Db 80 TCYEDQGISYRGTSWSTAESGAECTWNSSALAKPYSGRRPDALRLGLGNHNYCRNPDRD 139
 QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
 Db 140 SKPCYVFKAGYSSEFCSTPACSEG 165

RESULT 10
 Q86YK8 PRELIMINARY; PRT; 562 AA.
 ID Q86YK8
 AC Q86YK8;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Tissue plasminogen activator.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liu Y., Xu L., Zeng Y., He X.;
 RL "CDNA of tissue plasminogen activator";
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY221101; AAC34406.1; -.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibrinctn1.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00019; KRINGLE.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00058; FNI; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS00070; KRINGLE_2; 2.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 SQ SEQUENCE 562 AA; 62902 MW; 837D98392F5EDD1F CRC64;

Query Match 40.8%; Score 226; DB 4; Length 562;
 Best Local Similarity 47.7%; Pred. No. 4.1e-18;
 Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPDR 61
 Db 126 TCYEDQGISYRGTSWSTAESGAECTWNSSALAKPYSGRRPDALRLGLGNHNYCRNPDRD 185
 QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
 Db 186 SKPCYVFKAGYSSEFCSTPACSEG 211

RESULT 11
 Q8NG20 PRELIMINARY; PRT; 90 AA.
 ID Q8NG20
 AC Q8NG20;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Plasminogen/activator kringle.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

```

[1]
RN SEQUENCE FROM N.A.
RA Dou D.;
RP "Production of kringle fragment.";
RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF282882; AAM52248.1; -.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
SQ SEQUENCE 90 AA; 9804 MW; A33887F9DF4C7B1 CRC64;

Query Match 39.4%; Score 218.5; DB 4; Length 90;
Best Local Similarity 49.4%; Pred. No. 4.2e-18;
Matches 41; Conservative 5; Mismatches 36; Indels 1; Gaps 1;

QY 3 CYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNR 62
Db 8 CYFGNSAVRGTHSLTESGASCLPWSMILIGKVYTAQNPSAQLGLGKHNYCRNPDGDA 67
QY 63 RPYCVQVGLKPLVQSCMWHDC 85
Db 68 KPWCIYT-TNPRKLYDYCDVPQCA 89

RESULT 12
Q9BZW1 PRELIMINARY; PRT; 395 AA.
AC Q9BZW1
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Neocatal thrombolytic agent alpha-form (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF260825; AAK11956.1; -.
DR HSSP; P00750; 1PK2.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.

```

```

DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; 1.
DR PROSITE; PS0135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON TER 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 39.4%; Score 218.5; DB 4; Length 395;
Best Local Similarity 44.7%; Pred. No. 2.2e-17;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;

QY 3 CYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNR 62
Db 48 CYFGNSAVRGTHSLTESGASCLPWSMILIGKVYTAQNPSAQLGLGKHNYCRNPDGDA 107
QY 63 RPYCVQVGLKPLVQSCMWHDCAD-GKPSPP 95
Db 108 KPWCYHLNKRRLTWECYCDVPCSTCGLRQYSQ 141

RESULT 13
Q8SQ23 PRELIMINARY; PRT; 562 AA.
ID Q8SQ23
AC Q8SQ23
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Ding Y.; Xue J.; Bartlett J.D.;
RT "Plasminogen activator in tooth tissues";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AAM00297.1; -.
DR HSSP; P00761; 1AN1.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; 1.
DR PROSITE; PS0135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;

```

KW Serine protease.
SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101E8 CRC64;
Query Match 38.1%; Score 211; DB 6; Length 562;
Best Local Similarity 43.2%; Pred. No. 2.4e-16;
Matches 38; Conservative 10; Mismatches 40; Indels 0; Gaps 0;
QY 2 TCYEGNGHFVRGKASDTMGRPCLPWNSATVLQQTTHAHSRSDALQLGLGKHNYCRNPDR 61
DB 126 TCYEDQGITRGTSWTESGAEACWNWITSGLASPNYGRPRDPAVKLGLGNHNYCRNPDKD 185
QY 62 RRPWCYVQVGLKPLVQECMVHDCADGKK 89
DB 186 SKPWCYIFKAEKSPDFCSTACTKE 213
RESULT 14
Q8K0D2 PRELIMINARY; PRT; 517 AA.
AC Q8K0D2
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUES=Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC031775; AAH31775.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR00742; EGF_2.
DR InterPro; IPR001891; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TYP SPc; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS0240; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN HIS; 1.
DR PROSITE; PS00135; TRYPSIN SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;

Query Match 37.9%; Score 210; DB 11; Length 517;
Best Local Similarity 43.5%; Pred. No. 2.9e-16;
Matches 40; Conservative 14; Mismatches 34; Indels 4; Gaps 2;

QY 3 CYEGNGHFVRGKASDTMGRPCLPWNSATVLQQTTHAHSRSDALQLGLGKHNYCRNPDR 62
DB 151 CYVGGSYRGKVKTKVQNPCLYNWNSLLQLQTYMFMEDAETHGIAEHFRCRNPDDGH 210
QY 63 RRPWCYVQVGLKPLVQECMVHDCADGKPPSP 93
DB 211 KPWCFVKVNSEKVRKWCYCDTVC---PVPDTP 239
RESULT 15
Q800Y7 PRELIMINARY; PRT; 540 AA.
AC Q800Y7
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hepatocyte growth factor activator (Fragment).
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2122796; PubMed=12128063;
RA Holsberger D.R.; Becker A.E.; Thurston R.J.; Rice C.D.;
RT "Expression of a hepatocyte growth factor activator protein in turkey
RL (Meleagris gallopavo) deferent duct epithelial cells.";
RN [2]
RP SEQUENCE FROM N.A.
RA Holsberger D.R.; Becker A.E.; Thurston R.J.; Rice C.D.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; A216598; AAC46038.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR00742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctn.
DR InterPro; IPR00562; FN Type II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TYP SPc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRINECTIN 1; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS00240; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN HIS; 1.
DR PROSITE; PS00135; TRYPSIN SER; 1.
FT NON TER 1 540
FT NON TER 540
SQ SEQUENCE 540 AA; 61159 MW; 0BB3B4A89C0B577F CRC64;

Query Match 37.0%; Score 205; DB 13; Length 540;
Best Local Similarity 41.3%; Pred. No. 1.2e-15;
Matches 38; Conservative 12; Mismatches 40; Indels 2; Gaps 1;

```
QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSALQGLGKHNVCNPNR 62
Db 179 CYRGNGTEYRGAKTITISGSLPWSLILYRELHVDVSERKAVQLGGLGPFYCRNPDE 236
QY 63 RPWCYVQVGLKPLVQECMVHDCADGKPKSPSP 94
Db 239 KWCYIMKONSLSWEYCNITSCASRER--RPP 268

RESULT 16
Q14520 PRELIMINARY; PRT; 560 AA.
AC Q14520; 000663;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE HGF activator like protein (hyaluronan binding protein 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kitamura N.,
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96425001; PubMed=8827452;
RA Choi-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
RA Tomita M.;
RT "Purification and characterization of a novel hyaluronan-binding
RT protein (PHBP) from human plasma: it has three EGF, a kringle and a
RT serine protease domain, similar to hepatocyte growth factor
RT activator."
RL J. Biochem. 119:1157-1165(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; D49742; BA08576.1; -
DR EMBL; S83182; AAB46909.1; -
DR EMBL; BC031412; AAB31412.1; -
DR PIR; JC4795; JC4795.
DR HSSP; P00763; LDPO.
DR MEROPS; S01.033; -
DR Genew; HGNC:4798; HAP2.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0005539; F:glycosaminoglycan binding; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR008209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS50240; TRYPSIN_HIS; 1.
DR PROSITE; PS00134; TRYPSIN_SER; 1.
DR EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;
```

```
Query Match 36.0%; Score 199.5; DB 4; Length 560;
Best Local Similarity 40.2%; Pred. No. 5.5e-15;
Matches 39; Conservative 15; Mismatches 40; Indels 3; Gaps 2;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSALQGLGKHNVCNPNR 62
Db 194 CYVGDGYSYRGKVNRTVQHQACLYWNSHLLQENYNNFMEDAETHGIGHNFCRNPDA 253
QY 63 RPWCYVQVGLKPLVQECMVHDC--DGKKPSPSP 96
Db 254 KPCFKIKVTNDKVKWEYCDVSACSADQVAYPEESPT 290

RESULT 17
Q90675 PRELIMINARY; PRT; 202 AA.
AC Q90675;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Tissue-type plasminogen activator (Fragment).
GN TPA.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn;
RX MEDLINE=97195025; PubMed=9047000;
RA Johnson A.L., Bridgham J.T., Anthony R.V.;
RT "Expression of avian urokinase and tissue-type plasminogen activator
RT messenger ribonucleic acid during follicle development and atresia."
RL Biol. Reprod. 56:581-588(1997).
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; U31988; AAA74955.1; -
DR HSSP; P00750; IRTF.
DR MEROPS; S01.232; -
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 202
SQ SEQUENCE 202 AA; 23148 MW; 049DB342941D83AFC CRC64;

Query Match 34.5%; Score 191; DB 13; Length 202;
Best Local Similarity 43.9%; Pred. No. 1.8e-14;
Matches 36; Conservative 7; Mismatches 39; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSALQGLGKHNVCNPNR 62
Db 41 CYTGNGLAYRGTRSRKSGFSLPWNVFLTSKIYTALEEQRRALGLGKHNCRNP 100
QY 63 RPWCYVQVGLKPLVQECMVHDC 84
Db 101 QPWCHWKDRQLTWBYCDVPPQC 122
```

RESULT 18
Q8VCS4 PRELIMINARY; PRT; 653 AA.

AC Q8VCS4; PRELIMINARY; PRT; 653 AA.

AT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RC TISSUE=Liver;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; SC019376; AAH19376.1; -.

DR HSSP; P00761; IAN1.

DR GO; GO:0005576; C:extracellular; IEA.

DR GO; GO:0004263; F:chymotrypsin activity; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0004295; F:trypsin activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR009003; Cys_Ser_trypsin.

DR InterPro; IPR00742; EGF 2.

DR InterPro; IPR006209; EGF like.

DR InterPro; IPR000083; Fibrinctn1.

DR InterPro; IPR00562; FN Type II.

DR InterPro; IPR006210; EGF-like.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Peptidase_S1.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00039; fn1; 1.

DR Pfam; PF00040; fn2; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00013; FNTYPEII.

DR ProDom; PD000995; FN_Type_II; 1.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00181; EGF; 2.

DR SMART; SM00059; FN2; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00022; EGF_1; 2.

DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS01253; FIBRONECTIN_1; 1.

DR PROSITE; PS00023; FIBRONECTIN_2; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

DR PROSITE; PS00134; TRYPSIN_DOM; 1.

DR PROSITE; PS00240; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;

KW Kringle; Protease; Serine protease.

SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 33.9%; Score 186; DB 11; Length 653;

Best Local Similarity 53.0%; Pred. No. 1.5e-13;

Matches 35; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

QY 3 CYEGNGHYRGKASTDTMGRPCLPNWSATVLOQTYHAHRSDALQGLGKHNYCRNPDR 62

DB 283 CFLNGTETRGVASTAASGLSCLANNSDLLYQELHVDVAAAVILGLGPHAYCRNPDK 342

QY 63 RWCYV 68

DB 343 RWCYV 348

RESULT 19
O35727 PRELIMINARY; PRT; 597 AA.

AC O35727; PRELIMINARY; PRT; 597 AA.

AT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Factor XII.

GN F12.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Schloesser M.; Schwager S.; Engel W.;

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; X9571; CAA67891.1; -.

DR HSSP; P00760; IAO7.

DR MEROPS; S01.211; -.

DR MGD; MGI:1891012; F12.

DR GO; GO:0005576; C:extracellular; IEA.

DR GO; GO:0004263; F:chymotrypsin activity; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0004295; F:trypsin activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR009003; Cys_Ser_trypsin.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR000083; Fibrinctn1.

DR InterPro; IPR000562; FN_Type_II.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Peptidase_S1.

DR InterPro; IPR001314; Peptidase_S1A.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00039; fn1; 1.

DR Pfam; PF00040; fn2; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00013; FNTYPEII.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000995; FN_Type_II; 1.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00181; EGF; 2.

DR SMART; SM00059; FN2; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00022; EGF_1; 2.

DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS01253; FIBRONECTIN_1; 1.

DR PROSITE; PS00023; FIBRONECTIN_2; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

DR PROSITE; PS00240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;

KW Serine protease.

SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FEA CRC64;

Query Match 33.8%; Score 187.5; DB 11; Length 597;

Best Local Similarity 37.7%; Pred. No. 1.6e-13;

Matches 40; Conservative 10; Mismatches 39; Indels 17; Gaps 3;

QY 2 TCYEGNGHYRGKASTDTMGRPCLPNWSATVLOQTYHAHRSDALQGLGKHNYCRNP 60

DB 283 CFLNGTETRGVASTAASGLSCLANNSDLLYQELHVDVAAAVILGLGPHAYCRNPDK 342

QY 63 RWCYV 68

DB 343 RWCYV 348

	DB	216	TCTYEGRLSYRGQAGTOSGAPCRQM-----TVETATRNMTKQLSGLGHAAFCRNPDPN	271
	QY	61	RARPWCYVQGVLKLPLVQCWMHDC-----ADKKPSPPP	94
	DB	272	DTRPCWCFWSGDRLSDWDYCGLQCQTPTFAPLWPESQEERSQPAP	317

```

RESULT 20
QB0YCS
ID ID QB0YCS PRELIMINARY; PRT; 609 AA.
```

AC QB0YCS;

DT 01-JUN-2003 (TEMBLrel_24, Created)

DT 01-JUN-2003 (TEMBLrel_24, Last sequence update)

DT 01-OCT-2003 (TEMBLrel_25, Last annotation update)

DE Similar to coagulation factor XII (Hageman factor) (Fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_Taxid=10090;

[1]

RN SEQUENCE FROM N.A.

RP STRAIN=FVB/N; TISSUE=Liver;

RC Strausberg R.;
RA Submitted (RAR-2003) to the EMBL/GenBank/DDBJ databases.
RL ENBL; BC049867; AAH49867.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:tryptsin activity; IEA.
DR DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser_trypsin.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibnctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; LEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_Sl.
DR InterPro; IPR001314; Peptidase_SIA.
PFam; PF00008; EGF; 2.
PFam; PF00039; fn1; 1.
PFam; PF00040; fn2; 1.
PFam; PF00051; kringle1; 1.
PFam; PF00089; tryptsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00013; FNTYPEII.
PRINTS; PR00018; KRINGLE.
PrDom; PD000995; FN Type II; 1.
DrDom; PD000385; Kringle; 1.
SMART; SMC0181; EGF; 2.
SMART; SMC00179; EGF_CA; 2.
SMART; SMC00058; FN1; 1.
SMART; SMC00059; FN2; 1.
SMART; SMC0130; KR; 1.
SMART; SMC0020; TYPD_SPC; 1.
PROSITE; PS00022; EGF_1; 2.
PROSITE; PS01186; EGF_2; 1.
PROSITE; PS01253; FIBONECTIN_1; 1.
PROSITE; PS00023; FIBONECTIN_2; 1.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS00070; KRINGLE_2; 1.
PROSITE; PS50240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.

FT NON TER
SQ SEQUENCE 609 AA; 66783 MW; DF97D4B2369BD2 CRC64;

Query Match 33.8%; Score 187.5; DB 11; Length 609;
Best Local Similarity 37.7%; Pred. No. 1.ee-13;
Matches 40; Conservative 10; Mismatches 39; Indels 17; Gaps :

	QY	2	TCTYEGHLFRGKASTDGTMRCPLPNSATVLCQQTY-HAHRSDALQLGLKHNYCNRPN	60
--	----	---	---	----

Dd		228	TCTGEGSLYSRGQAGTTQSAPQRW----	TVEATYRNMTKQAQLSWGLGHHAFCF
Qy		61	RERPWCYYVQNGKLPLVGECMWHD-----ADCKKSPP	94
Dd		284	DTRPNCFWSGDRLLSDWDYCGLQCQTPTFAPLVPESQBSQP	329
RESULT	21			
O97507	PRELIMINARY;	PRT;	616 AA.	
ID	O97507			
AC	O97507; TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DS	FALL.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus,			
CX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RA	Takahashi T., Kihara T.;			
RT	"Porcine liver factor XII.";			
RL	Submitted (JAN-1999) to the EMBL/genBank/DBJ databases.			
CC	-I SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-I SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.			
EMBL:	A8022426; BAA37148.1; --			
HSP:	P00763; IDPO.			
KR:	--			
GO:	GO: 0005576; C:extracellular; IEA.			
GO:	GO: 0004263; F:chymotrypsin activity; IEA.			
GO:	GO: 0008233; F:peptidase activity; IEA.			
GO:	GO: 0004295; F:tryptic activity; IEA.			
GO:	GO: 0006508; P:proteolysis and peptidolysis; IEA.			
InterPro:	IPIR009003; Cys Ser_trypsin.			
InterPro:	IPIR00742; EGf_2			
InterPro:	IPIR006209; EGf-like.			
InterPro:	IPIR000083; Fibronctn.I.			
InterPro:	IPIR000562; FN_Type-II.			
InterPro:	IPIR006210; IEGF.			
InterPro:	IPIR000001; Kringle.			
InterPro:	IPIR001254; Peptide_S1.			
InterPro:	IPIR001314; Peptide_SIA.			
Pfam:	PF00008; EGF_2.			
Pfam:	PF00039; fnl_1.			
Pfam:	PF00040; fn2_1.			
Pfam:	PF00051; kringle_1.			
Pfam:	PF00089; trypsin_1.			
PRINTS:	PR00722; CHYMOTRYPSIN.			
PRINTS:	PRO0013; FNTYPEII.			
PRINTS:	PRO0019; KRINGLE.			
ProDom:	PD000995; FN_Type-II; 1.			
ProDom:	PD000395; Kringle; 1.			
SMART:	SMO0181; EGF_2.			
SMART:	SMO0058; FN1; 1.			
SMART:	SN00059; FN2; 1.			
SMART:	SMO0130; KP; 1.			
SMART:	SMO0020; Tryp_Spc; 1.			
PROSITE:	PS00022; EGF_1; 2.			
PROSITE:	PS01186; EGF_2; 1.			
PROSITE:	FS01253; FIBRONECTIN_1; 1.			
PROSITE:	FS00023; FIBRONECTIN_2; 1.			
PROSITE:	PS00021; KRINGLE_1; 1.			
PROSITE:	PS00070; KRINGLE_2; 1.			
PROSITE:	PS00240; TRYPSIN_DOM; 1.			
PROSITE:	PS00134; TRYPSIN_HIS; 1.			
PROSITE:	PS00135; TRYP SIN SER; 1.			
KW	EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;			
KW	Serine protease.			
SEQ	SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBDIA9 CRC64;			

```

Best Local Similarity 41.2%; Pred. No. 1.9e-12;
Matches 42; Conservative 7; Mismatches 40; Indels 13; Gaps 4;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQY-HAHRSDALQLGLGKHNCRPNDR 61
DB 217 CYSRGLSYRGMAQTTLTGAPCPWAS-----BATYNNMTAEQALNWLGLGDFACRPNPD 272
QY 62 RRPWCYVQVGLRPLVQECMVHDCAD--GKKP-----SSPPE 95
DB 273 TRPWCFFWGRGDLQSWOYCLARCOAPIGEAPPILFTQSPSE 314

RESULT 22
Q8AXX3 PRELIMINARY; PRT; 421 AA.
AC Q8AXX3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Kremen2.
GN Kremen2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Fipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-22308973; PubMed-12421700;
RA Davidson G., Mao B., Del Barco Barrantes I., Niehrs C.;
RT "Kremen proteins interact with Dkk1 to regulate anteroposterior
RT CNS patterning.";
RL Development 129:5587-5596(2002).
DR EMBL; AV150813; AAN64661.1; -.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; KRINGLE; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00321; WSC; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE; 1.
DR PROSITE; PS50070; KRINGLE; 2; 1.
SQ SEQUENCE 421 AA; 46274 MW; 74D04DB1682CD15F CRC64;

Query Match 31.1%; Score 172.5; DB 13; Length 421;
Best Local Similarity 50.7%; Pred. No. 6.2e-12;
Matches 34; Conservative 4; Mismatches 26; Indels 3; Gaps 2;

QY 3 CYEGNGHFYRGKAS-TDTMGRPCLPWNSATVLQOQY-HAHRSDALQLGLGKHNCRPNDR 61
DB 29 CFTVNGRDYRGTVSQAGPETPCLYNQTT--QHLNQAQSDPDGELGLGKHNCRPNDR 86
QY 62 RRPWCYV 68
DB 87 VQPCYV 93

RESULT 23
Q8IZZ5 PRELIMINARY; PRT; 615 AA.
AC Q8IZZ5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Coagulation factor XII-Mie.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wada H., Nishiooka J., Nakatani K., Kasai Y., Abe Y., Nobori T.;
RT "Molecular characterization of coagulation factor XII-Mie.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095845; BAC23095.1; -.
DR GO; GO:000576; C:extracellular; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006203; EGF_like.
DR InterPro; IPR000083; Fibrinctni.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 30.8%; Score 170.5; DB 4; Length 615;
Best Local Similarity 48.5%; Pred. No. 1.6e-11;
Matches 33; Conservative 6; Mismatches 24; Indels 5; Gaps 2;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQY-HAHRSDALQLGLGKHNCRPNDR 60
DB 216 SCYDGRGLSYRGLARTTLTGAPCPWAS-----BATYNNMTAEQALNWLGLGDFACRPNDR 271
QY 61 RRPWCYV 68
DB 272 DIRPWCYV 279

RESULT 24
Q8AXY6 PRELIMINARY; PRT; 947 AA.
AC Q8AXY6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Muscle-specific receptor tyrosine kinase Musk.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

```


RA Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation is Associated with Deletion of a Single Exon in a Null
RT Allele";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF029691; AAB97886.1; -
DR HSSP; P00747; 2PK4.
DR MEROPS; S01.999; -
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004235; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;
KW Serine protease.
FT NON_TER 1
SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;
Query Match 28.5%; Score 158; DB 6; Length 454;
Best Local Similarity 39.8%; Pred. No. 3.5e-10;
Matches 35; Conservative 10; Mismatches 31; Indels 12; Gaps 4;
QY 3 CYEGNGHFYRGKASTDTMGPRCLPWSATVQ--OTVHAHRSALQLGLGKHNYCRNP 60
DB 102 CYHGGQSVRGSTFTVTRGTCQSSMTPHQKETPENHNDGLTW-----NYCRPDA 156
QY 61 RRRPWCYVQVGLKPLV--QECMVHDCAD 86
DB 157 DTGPNWCF--MDPSRWYECNLTGRCSD 181
RESULT 27
Q9UIR5 PRELIMINARY; PRT; 113 AA.
AC Q9UIR5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2003 (TREMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21181705; PubMed=11285247;
RA Ogorekova M., Kraft H.G., Enholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect lip(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158663; AAF03680.1; -
DR EMBL; AF158662; AAF03680.1; JOINED.
DR HSSP; P00747; 1PMK.
DR InterPro; IPR000001; Kringle.

DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12685 MW; F3D65681D9B5253A CRC64;
Query Match 28.3%; Score 157; DB 4; Length 113;
Best Local Similarity 34.8%; Pred. No. 9.9e-11;
Matches 39; Conservative 13; Mismatches 36; Indels 24; Gaps 6;
QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWSATV--LQOTVHAHRSALQLGLGKHNYCRNP 58
DB 9 RCYHNGQSVRGSTFTVTRGTCQSSMTPHQKETPENHNDGLTW-----NYCRNP 63
QY 59 DNRERPWCYVQVGLKPLV--QECMVHDCADG-----KKPS-SPPEE 96
DB 64 DADTGPWCFT--MDPSRWYECNLTGRCSDTGTGTVVAPPTVIQVPSLGPSE 112
RESULT 28
O18783 PRELIMINARY; PRT; 806 AA.
AC O18783;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Plasminogen.
OS Macropus eugenii (Tamar wallaby).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98004511; PubMed=9342350;
RA Lawn R.M., Schwartz K., Patthy L.;
RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; AF012297; AAB65760.1; -
DR HSSP; P00747; 5HPG.
DR MEROPS; S01.233; -
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0004235; F:trypsin activity; IEA.
DR GO; GO:0007596; F:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.


```
Qy 60 NRRR-PWCYVQVGLKPLVQECMVHDC--DGKK 89
Db 370 GQMEGFWCFQ-QKNVRVELCDVPPCSPRDGSK 401

RESULT 31
Q8BSP6 PRELIMINARY; PRT; 944 AA.
ID QBSP6
AC QBSP6
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Receptor tyrosine kinase-like orphan receptor 2.
GN ROR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=22354693; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK031112; BAC27258.1; --
DR MGI; MGI:1347521; Ror2.
DR GO; GO:0001501; P:skeletal development; IMP.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG c2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR01245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TyrK; 1.
DR PROSITE; PS50038; Fz; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS50070; KRINGLE 2; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
SQ SEQUENCE 944 AA; 105053 MW; 1BBD416BE3170401 CRC64;

Query Match 27.7%; Score 153.5; DB 11; Length 944;
Best Local Similarity 41.9%; Pred. No. 2.7e-09;
Matches 39; Conservative 9; Mismatches 32; Indels 13; Gaps 7;

Qy 3 CYEGNGHYRGKASTDTGRCPLPWSATVLCQTYHAHR---SDALQGLGKHNYCRNPD 59
Db 316 CYNGSGADYRGVASTTKSGHQCPW-----ALQHP-HSHRLSSTFPPELG-GGHAYCRNPG 369

Qy 60 NRRR-PWCYVQVGLKPLVQECMVHDC--DGKK 89
Db 370 GQMEGFWCFQ-QKNVRVELCDVPPCSPRDGSK 401

RESULT 32
Q9UIR7 PRELIMINARY; PRT; 113 AA.
ID Q9UIR7
AC Q9UIR7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21181705; PubMed=11285247;
RA Gurelko M., Kraft H.G., Ehnmolm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158659; AAF03678.1; JOINED.
DR HSSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR GlycoProtein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12815 MW; 4F50ADF8708548CB CRC64;

Query Match 27.6%; Score 153; DB 4; Length 113;
Best Local Similarity 39.3%; Pred. No. 2.9e-10;
Matches 33; Conservative 8; Mismatches 35; Indels 8; Gaps 3;

Qy 3 CYEGNGHYRGKASTDTGRCPLPWSATVLCQTYHAHRSDALQGLGKHNYCRNPDNR 62
Db 11 CYHGQSYRGSGFSTVTGRTQSHSMT---PHGHQRTTEYYPNGGLTRNYCRNPD 67

Qy 63 RPWCYVQVGLKPLV--QECMVHDC 84
Db 68 RPWCYT---MDPSVRWEYCNLTRC 88

RESULT 33
Q8WNR1 PRELIMINARY; PRT; 359 AA.
ID Q8WNR1
AC Q8WNR1
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Pirie-Shepherd S.R., Coffman K.T., Resnick D., Chan R., Kisker O.,
RA Folkman J., Waters D.J.;
RT "Angiostatin is Detectable in the Urine of Dogs with Spontaneous Bone
RT Cancer.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; AY069985; AAL58519.1; --
DR GO; GO:0005509; F:calcium ion binding; IEA.
```

```

DR GO: GO:0003809; F:thrombin activity; IEA.
DR GO: GO:0007596; P:blood coagulation; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003966; Peptidase_S1A_pr.
DR Pfam: PF00051; kringle; 4.
DR PRINTS; PR0018; KRINGLE.
DR PRODOM; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR Glycoprotein; KRingle.
FT NON_TER 1
FT NON_TER 359
SQ SEQUENCE 359 AA; 41172 MW; 776035F4AB0BDD9E CRC64;

Query Match 27.3%; Score 151.5; DB 6; Length 359;
Best Local Similarity 44.3%; Pred. No. 1.6e-09;
Matches 31; Conservative 5; Mismatches 21; Indels 13; Gaps 3;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQ-----OLGLGKHYCNRPD 57
DQ 282 CYHNGSYAGTSTITGKQCQSWMT-----PHREKTPHFPEAGL-TWNYCRN 333
QY 58 PDNRPPWCY 67
DQ 334 PDADKSPWCY 343

RESULT 34
Q9ROW3 Q9ROW3 PRELIMINARY; PRT; 812 AA.
AC Q9ROW3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
GN PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RC MEDLINE=91250378; PubMed=1645711;
RA Kanakas J.J., Makker S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RL receptor site for plasminogen.";
RC J. Biol. Chem. 266:10825-10829 (1991).
CC -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; AJ242649; CAB46014.1; -.
DR HSSP; P00747; 1PMK.
DR GO: GO:0005576; C:extracellular; IEA.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR GO: GO:0004263; F:chymotrypsin activity; IEA.
DR GO: GO:0005179; F:hormone activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0004283; F:plasmin activity; IEA.
DR GO: GO:0003809; F:thrombin activity; IEA.
DR GO: GO:0004295; F:trypsin activity; IEA.
DR GO: GO:0007596; P:blood coagulation; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys_Ser_trypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQ-----OLGLGKHYCNRPD 59
DQ 376 CYQNGKSYRGTSSTITGKQCQSW-----VSMTPHSHSKTANFPDAGL-EMNYCRNP 429
QY 60 N-RRPWCYVQGLKPLV--QECMVHDCAD-----GKKPSP 93
DQ 430 NDQRPWCFT---TDPVRYEYCNLCRCSETGGVAESAIVQVPSAP 474

RESULT 35
Q91691 Q91691 PRELIMINARY; PRT; 716 AA.
AC Q91691;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Growth factor Livertine.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN RN SEQUENCE FROM N.A.
RP Ruiz i Altaba A., Thery C.;
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; U57455; AAB52574.1; -.
DR HSSP; P00747; 1CEA.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR GO: GO:0004263; F:chymotrypsin activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0003809; F:thrombin activity; IEA.
DR GO: GO:0004295; F:trypsin activity; IEA.
DR GO: GO:0007596; P:blood coagulation; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys_Ser_trypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003809; Pan_app.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR InterPro: IPR003966; Peptidase_S1A_pr.

```



```

Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
-1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
EMBL; BC010551; AAH10551.1; -.
HSP; P00761; IANI.
MGI; MGI:96080; Matl.
GO; GO:0007566; Piembryo implantation; IC.
InterPro; IPR009003; Cys_Sek_trypsin.
InterPro; IPR000001; Kringle.
InterPro; IPR030314; PAN.
InterPro; IPR003609; Pan_app.
InterPro; IPR001254; Peptidase_S1.
InterPro; IPR001314; Peptidase_S1A.
InterPro; IPR003966; Peptidase_S1A_pr.
Pfam; PF00051; kringle; 4.
Pfam; PF00024; PAN; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; KRINGLE.
PRINTS; PR01505; PROTHROMBIN.
ProDom; PD000395; Kringle; 4.
SMART; SM001130; KR; 4.
SMART; SM00473; PAN_AP; 1.
SMART; SM00020; TRY_P_SPC; 1.
PROSITE; PS00021; KRINGLE 1; 4.
PROSITE; PS00070; KRINGLE 2; 4.
PROSITE; PS00240; TRYPSIN_DOM; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 716 AA; 80693 MW; 12474C48A7D4B46D CRC64;

Query Match 26.8%; Score 147.5; DB 11; Length 716;
Best Local Similarity 36.5%; Pred. No. 1e-08;
Matches 38; Conservative 11; Mismatches 34; Indels 21; Gaps 7;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQYTHAHR-----SDALQLGLGHNYCRN 57
Db 379 CYHSGSEQRGVSVKTRKGVCQHNSET-----PHKQPTFTSAPQAGL-EANFCRN 430

QY 58 PD-NRRRPWCYGVGLKP--LVQECMVHDCADGKKFS--SPPEE 96
Db 431 PDGDSHGPMWCYT---LDPDILFDYCALQRCDDDPSPILDPPDQ 471

RESULT 40
Q9UIR6 PRELIMINARY; PET; 113 AA.
AC Q9UIR6; 2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21181705; PubMed=11285247;
RA Ogarekova M., Kraft H.G., Enholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158861; AAF03679.1; -.
DR EMBL; AF158660; AAF03679.1; JOINED.
DR HSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.

```

```

FT CHAIN 20 810 PLASMINOGEN.
SQ SEQUENCE 810 AA; 90555 MW; B05C7D74B0D020B3C CRC64;

Query Match 26.4%; Score 146.5; DB 4; Length 810;
Best Local Similarity 37.4%; Pred. No. 1.5e-08;
Matches 37; Conservative 13; Mismatches 32; Indels 17; Gaps 7

QY 3 CYEENGHFRYKASTDTMGPRCLFPWNSATVLCQTYHAHR---SDALQLGLGKHNYCRNPD 59
DB 377 CYHGQGSYRTSITTTGKKCSWS-----MTPRHQKTPENYPNAGL-TWNYCENPD 430
QY 60 NRRRPMCVTVGLKPLV--QECMVHDCADKKFS--SPP 94
DB 431 ADKGWCFP---TDPFVRWEYCNLKKS-GTEASVWAPP 465

RESULT 42
Q9BBR6 PRELIMINARY; PRT; 393 AA.
AC Q9BBR6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Similar to receptor tyrosine kinase-like orphan receptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC006374; AAH06374.1; -.
DR HSSP; P00747; 1CEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00408; Igc2; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Immunoglobulin domain; Kinase; Kringle; Receptor.
SQ SEQUENCE 393 AA; 43825 MW; 1F93DCBBBF53855 CRC64;

Query Match 26.4%; Score 146; DB 4; Length 393;
Best Local Similarity 44.1%; Pred. No. 7.8e-09;
Matches 30; Conservative 5; Mismatches 25; Indels 8; Gaps 3

QY 3 CYEENGHFRYKASTDTMGPRCLFPWNSATVLCQTYHAHRSDALQLG--LGKHNYCRNPD 60
DB 313 CYNSTGVDRGTVSVTKSGRCQPWNS-----QYPHTTFTALRFPFLNGHSGYCRNPGN 367
QY 61 RRR-PWCY 67
DB 368 QKEAPWCF 375

RESULT 43
Q8BNP9

```

```

ID Q8BNP9 PRELIMINARY; PRT; 937 AA.
AC Q8BNP9;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Receptor tyrosine kinase-like orphan receptor 1.
GN ROR1
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adipose tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK080896; BAC38068.1; -.
DR MGD; MGI:1347520; Ror1.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
SQ SEQUENCE 937 AA; 104118 MW; CB440323CA66EC28 CRC64;

Query Match 26.4%; Score 146; DB 11; Length 937;
Best Local Similarity 44.1%; Pred. No. 2e-08;
Matches 30; Conservative 5; Mismatches 25; Indels 8; Gaps 3;

QY 3 CYEGNGHYRGKASDTMTGRPCLPWN SATVLQQTTHAHRSDALQLG--LGKHNYCNP DN 60
DB 313 CYNSTGVYRGTVSVTKSGRCQCPWNS-----QYPHTHSFTALRFPPLNGHGHSCYCNPGN 367

QY 61 RRR-PWCY 67
DB 368 QKEAPWCF 375

RESULT 44
```

```

Q8BG10 PRELIMINARY; PRT; 937 AA.
ID Q8BG10;
AC Q8BG10;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Receptor tyrosine kinase-like orphan receptor 1.
GN ROR1
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adipose tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK046699; BAC32840.1; -.
DR MGD; MGI:1347520; Ror1.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
SQ SEQUENCE 937 AA; 104088 MW; D6F2D84B67D03C69 CRC64;

Query Match 26.4%; Score 146; DB 11; Length 937;
Best Local Similarity 44.1%; Pred. No. 2e-08;
Matches 30; Conservative 5; Mismatches 25; Indels 8; Gaps 3;

QY 3 CYEGNGHYRGKASDTMTGRPCLPWN SATVLQQTTHAHRSDALQLG--LGKHNYCNP DN 60
DB 313 CYNSTGVYRGTVSVTKSGRCQCPWNS-----QYPHTHSFTALRFPPLNGHGHSCYCNPGN 367

QY 61 RRR-PWCY 67
DB 368 QKEAPWCF 375
```


OC Rhabdittidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
ON (1)
RP SEQUENCE FROM N.A.
RA Berks M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."; Science 282:2012-2018(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Northmore B.J.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=N2;
RA Koga M., Takeuchi M., Tameishi T., Ohshima Y.;
RT "Control of DA1-7 TGF expression and neuronal process development by a
RT receptor tyrosine kinase KIN-8 in C. elegans.";
RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; Z35595; CAA84639.2; -;
DR EMBL; Z47808; CAA84639.2; JOINED.
DR EMBL; Z47808; CAB61003.1; -;
DR EMBL; Z35595; CAB61003.1; JOINED.
DR EMBL; AJ132947; CAC29085.1; -;
DR PIR; T18840; T18840.
DR HSSP; P11362; IFGK.
DR WormPep; C01G6.8b; CE24774.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR035398; IgC2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF000047; ig; 1.
DR Pfam; PF000051; kringle; 1.
DR Pfam; PF00089; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; FALSE_NEG.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR KEGG; ATP-binding; Immunoglobulin domain; Kinase; Kringle; Transferase;
DR KEGG; Tyrosine-protein kinase.
SQ SEQUENCE 902 AA; 101038 MW; 2A03D76D07C552B5 CRC64;

Query Match 26.2%; Score 145; DB 5; Length 902;
Best Local Similarity 35.2%; Pred. No. 2.Se-08;
Matches 32; Conservative 12; Mismatches 35; Indels 12; Gaps 5;
OV 3 CYEGNGHFYRGKASTDTMGPCPLPWSATVLOOTYHAHRSDALQLGKGNCRPNPDR 62

Db 328 CYNVSGTQYEGTVAQTSKQCAPWIDST--SRDFNVHRPPEL---MNSKNYCRNPGGKK 382
Qy 63 -RPMCYVQVGLKPLVOE--CWVHDCADGKKP 90
Db 383 SRWCY-----SKPMQGEYCDVPQCPSPDMYP 409

RESULT 47
Q9BLX1
ID Q9BLX1 PRELIMINARY; PRT; 928 AA.
AC Q9BLX1
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Receptor tyrosine kinase (C. elegans Cam-1 protein) (corresponding
sequence C01G6.8a)
DE sequence C01G6.8a
GN C01G6.8 OR KIN-8 OR CAM-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;
OC Rhabditidae; Pelodirinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN (1)
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RA Koga M.; Take-uchi M.; Tameishi T.; Ohshima Y.;
RT "Control of DAF-7 TGF expression and neuronal process development by a
RT receptor tyrosine kinase KIN-8 in C. elegans.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RA Bekas M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN (4)
RP SEQUENCE FROM N.A.
RA Mortimore B.J.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AJ132946; CAC29084.1; -;
DR EMBL; Z35595; CAD36478.1; -;
DR EMBL; Z47808; CAD36478.1; JOINED.
DR EMBL; Z47808; CAD36487.1; -;
DR EMBL; Z35595; CAD36487.1; JOINED.
DR HSP; P11362; IFGK.
DR WormPep; C01G6.8a; CE32563.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004740; P:transferase activity; IEA.
DR GO; GO:0004888; P:transmembrane receptor activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000577; FGGY kin.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR001245; Tyr pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00047; Ig_1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.

ProDom; PD000395; Kringle; 1.
ProDom; PD000001; Prot kinase; 1.
SMART; SM00408; IGC2; 1.
SMART; SM00130; KR; 1.
SMART; SM00219; TyKc; 1.
PROSITE; PS00933; FGGY_KINASES_1; 1.
PROSITE; PS50038; FZ; 1.
PROSITE; PS50835; IG_LIKE; 1.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS00070; KRINGLE_2; 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; immunoglobulin domain; Kinase; Kringle; Transferase;
KW Tyrosine-protein kinase.
SQ SEQUENCE 928 AA; 103864 MW; F13B8C9BCAB30D20 CRC64;
Query Match 26.2%; Score 145; DB 5; Length 928;
Best Local Similarity 35.2%; Pred. No. 2.6e-08;
Matches 32; Conservative 12; Mismatches 35; Indels 12; Gaps 5;
Qy 3 CYEGNHGYRKASTDTMGRLPCLPWSNATVQQYTHAHRSDALQGLGKNYCRNPNRR 62
Db 354 CYNVSGTQYEGTVAQTSKQCAPWIDST--SRDFNVHRPPEL---MNSKNYCRNPGGKK 408
Qy 63 -RPMCYVQVGLKPLVOE--CWVHDCADGKKP 90
Db 409 SRWCY-----SKPMQGEYCDVPQCPSPDMYP 435

RESULT 48
Q90ZN6
ID Q90ZN6 PRELIMINARY; PRT; 709 AA.
AC Q90ZN6;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hepatocyte growth factor-like 1.
GN MST1 OR HGFL1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN (1)
RN SEQUENCE FROM N.A.
RA Bassett D.I.; Wilson S.W.;
RT "Early expression of zebrafish Hepatocyte Growth Factor-Like 1
RT suggests a conserved role in vertebrate neural induction.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; AF370035; AAK54207.1; -;
DR HSP; P00761; 1AN1.
DR ZFIN; ZDB-GENE-020806-3; mst1.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; P:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.

Search completed: May 25, 2004, 14:57:19
Job time : 16.9018 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: May 25, 2004, 14:43:00 ; Search time 23.4741 seconds
(without alignments)
1155.508 Million cell updates/sec

Title: US-09-880-503-9
Perfect score: 554
Sequence: 1 KTCVEGNHFGYRGKASTDTM.....QECMVHDCADCKPSSPPEE 96

Scoring table: BLOSUM62
Gapop.10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	554	100.0	96	AAE16550	Human uPA
2	554	100.0	143	AAE16549	Human uPA
3	554	100.0	337	AAE75492	Human col
4	554	100.0	337	ABP41795	Human ova
5	554	100.0	365	AAE68854	Delta 1-4
6	554	100.0	378	AAW13635	Human pro
7	554	100.0	386	AAE66266	Bifunctio
8	554	100.0	389	AAW13636	Human pro
9	554	100.0	390	AAE66245	Bifunctio
10	554	100.0	390	AAE66247	Bifunctio
11	554	100.0	392	AAE66260	Bifunctio
12	554	100.0	392	AAE66264	Bifunctio
13	554	100.0	392	AAE66255	Bifunctio
14	554	100.0	392	AAE66259	Bifunctio
15	554	100.0	392	AAE66261	Bifunctio
16	554	100.0	392	AAE66258	Bifunctio
17	554	100.0	392	AAE66263	Bifunctio
18	554	100.0	392	AAE66254	Bifunctio
19	554	100.0	392	AAE66256	Bifunctio
20	554	100.0	392	AAE66257	Bifunctio
21	554	100.0	393	AAE66251	Bifunctio
22	554	100.0	393	AAE66253	Bifunctio
23	554	100.0	393	AAE66249	Bifunctio
24	554	100.0	393	AAE66252	Bifunctio
25	554	100.0	393	AAE66244	Bifunctio

26	554	100.0	393	2	AAE66250	Bifunctio
27	554	100.0	393	2	AAE95597	Chimeric
28	554	100.0	393	2	AAE95596	Chimeric
29	554	100.0	393	2	AAE99885	M36: fibr
30	554	100.0	395	2	AAE47902	Pro-uroki
31	554	100.0	395	2	AAE66265	Bifunctio
32	554	100.0	395	2	AAE66262	Bifunctio
33	554	100.0	396	2	AAE66246	Bifunctio
34	554	100.0	397	2	AAE66248	Bifunctio
35	554	100.0	401	2	AAW13637	Human pro
36	554	100.0	410	2	AAE23794	Prourokin
37	554	100.0	411	1	AAE50871	Sequence
38	554	100.0	411	2	AAE06244	Urokinase
39	554	100.0	411	2	AAE07904	Human pro
40	554	100.0	411	2	AAE07902	Human pro
41	554	100.0	411	2	AAE07903	Human pro
42	554	100.0	411	2	AAE05117	UK-S3 as
43	554	100.0	411	2	AAW13634	Human nat
44	554	100.0	411	2	AAE10057	Pro-uroki
45	554	100.0	411	2	AAE10058	Pro-uroki
46	554	100.0	411	2	AAE47956	PUK G16A
47	554	100.0	411	2	AAE47966	PUK G16A
48	554	100.0	411	2	AAE47962	PUK G38A
49	554	100.0	411	2	AAE40225	PUK 2/19
50	554	100.0	411	2	AAE47965	PUK N32P
51	554	100.0	411	2	AAE47958	PUK G38A
52	554	100.0	411	2	AAE47961	PUK G38A
53	554	100.0	411	2	AAE47959	PUK G32P
54	554	100.0	411	2	AAE47960	PUK G32P
55	554	100.0	411	2	AAE47964	PUK N32P
56	554	100.0	411	2	AAE47957	PUK N32P
57	554	100.0	411	2	AAE62997	Pro-uroki
58	554	100.0	411	2	AAE63008	Pro-uroki
59	554	100.0	411	2	AAE62999	Pro-uroki
60	554	100.0	411	2	AAE62994	Pro-uroki
61	554	100.0	411	2	AAE63003	Pro-uroki
62	554	100.0	411	2	AAE63001	Pro-uroki
63	554	100.0	411	2	AAE63006	Pro-uroki
64	554	100.0	411	2	AAE62992	Pro-uroki
65	554	100.0	411	2	AAE63002	Pro-uroki
66	554	100.0	411	2	AAE63007	Pro-uroki
67	554	100.0	411	2	AAE62995	Pro-uroki
68	554	100.0	411	2	AAE63000	Pro-uroki
69	554	100.0	411	2	AAE63004	Pro-uroki
70	554	100.0	411	2	AAE62996	Pro-uroki
71	554	100.0	411	2	AAE63005	Pro-uroki
72	554	100.0	411	2	AAE62991	Pro-uroki
73	554	100.0	411	2	AAE62998	Pro-uroki
74	554	100.0	411	2	AAE62993	Pro-uroki
75	554	100.0	411	2	AAE62926	Pro-uroki

ALIGNMENTS

RESULT 1
AAE16550
ID AAE16550 standard; protein; 96 AA.
XX
XX AAE16550;
AC
XX
XX
DT 09-APR-2002 (first entry)
XX
XX Human uPA kringle and connecting peptide.
XX
XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension; stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder; microvascular occlusion; angiotensin disorder; pulmonary fibrosis; asthma; tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing; clotting disorder; uterine contraction disorder; respiratory disease; adult respiratory distress syndrome; male impotence.

OS Homo sapiens.

XX WO200197752-A2.
 XX 27-DEC-2001.
 XX 13-JUN-2001; 2001WO-US018976.
 XX 20-JUN-2000; 2000US-0212874P.
 XX (UYPE-) UNIV PENNSYLVANIA.
 XX Cines DB, Higazi AA;
 XX WPI; 2002-122240/16.
 XX N-PSDB; AAD27083.
 XX Composition for modulating muscle cell and tissue contractility for
 XX treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 XX comprising domains from urokinase-type plasminogen activator.
 XX Claim 25; Fig 11; 117pp; English.
 XX The invention relates to a composition comprising one or more domains of
 XX urokinase-type plasminogen activator (uPA). The composition is used to
 XX modulate the contractility and angiogenic activity of a mammalian muscle,
 XX endothelial cell or tissue. The composition is used for treating stroke,
 XX hypotension, hypertension, atherosclerosis, heart attack, microvascular
 XX occlusions, thrombotic microangiopathies, surgically induced thrombotic
 XX disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 XX invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 XX diabetic retinopathy, wound healing, clotting disorder, uterine
 XX contraction disorder, male impotence, respiratory disease or condition
 XX such as asthma, adult respiratory distress syndrome, primary pulmonary
 XX hypertension, microvascular thrombotic occlusion, and a disorder
 XX associated with chronic intrapulmonary fibrin formation. The present
 XX sequence is human urokinase-type plasminogen activator (uPA) kringle and
 XX connecting peptide
 XX Sequence 96 AA;
 XX
 XX Query Match 100.0%; Score 554; DB 5; Length 96;
 XX Best Local Similarity 100.0%; Pred. No. 5.1e-45;
 XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 KTCYEGNGHFRGKASTDTMGRLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPDN 60
 XX Db 1 KTCYEGNGHFRGKASTDTMGRLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPDN 60
 XX
 XX QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
 XX Db 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
 XX
 XX RESULT 2
 XX AAE16549
 XX ID AAE16549 standard; protein; 143 AA.
 XX AC AAE16549;
 XX DT 09-APR-2002 (first entry)
 XX DE Human uPA amino terminal fragment (ATF) and connecting peptide.
 XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
 XX stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
 XX microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
 XX tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
 XX clotting disorder; uterine contraction disorder; respiratory disease;
 XX adult respiratory distress syndrome; amino terminal fragment; ATF;
 XX male impotence.
 XX Homo sapiens.
 XX OS
 XX

PN WO200197752-A2.
 PD 27-DEC-2001.
 PF 13-JUN-2001; 2001WO-US018976.
 PR 20-JUN-2000; 2000US-0212874P.
 PA (UYPE-) UNIV PENNSYLVANIA.
 PI Cines DB, Higazi AA;
 XX WPI; 2002-122240/16.
 DR N-PSDB; AAD27082.
 XX Composition for modulating muscle cell and tissue contractility for
 XX treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 XX comprising domains from urokinase-type plasminogen activator.
 XX Claim 24; Fig 1H; 117pp; English.
 XX The invention relates to a composition comprising one or more domains of
 XX urokinase-type plasminogen activator (uPA). The composition is used to
 XX modulate the contractility and angiogenic activity of a mammalian muscle,
 XX endothelial cell or tissue. The composition is used for treating stroke,
 XX hypotension, hypertension, atherosclerosis, heart attack, microvascular
 XX occlusions, thrombotic microangiopathies, surgically induced thrombotic
 XX disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 XX invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 XX diabetic retinopathy, wound healing, clotting disorder, uterine
 XX contraction disorder, male impotence, respiratory disease or condition
 XX such as asthma, adult respiratory distress syndrome, primary pulmonary
 XX hypertension, microvascular thrombotic occlusion, and a disorder
 XX associated with chronic intrapulmonary fibrin formation. The present
 XX sequence is human urokinase-type plasminogen activator (uPA) amino
 XX terminal fragment (ATF) and connecting peptide
 XX Sequence 143 AA;
 XX
 XX Query Match 100.0%; Score 554; DB 5; Length 143;
 XX Best Local Similarity 100.0%; Pred. No. 7.5e-45;
 XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 KTCYEGNGHFRGKASTDTMGRLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPDN 60
 XX Db 48 KTCYEGNGHFRGKASTDTMGRLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPDN 107
 XX
 XX QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
 XX Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 143
 XX
 XX RESULT 3
 XX AAG75492
 XX ID AAG75492 standard; protein; 337 AA.
 XX AC AAG75492;
 XX DT 03-SEP-2001 (first entry)
 XX DE Human colon cancer antigen protein SEQ ID NO:6256.
 XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 XX colorectal carcinoma; chromosome 10.
 XX Homo sapiens.
 XX OS
 XX WO200122920-A2.
 XX 05-APR-2001.
 XX 28-SEP-2000; 2000WO-US026524.
 XX

DT 25-MAR-2003 (revised)
 DT 22-NOV-1995 (first entry)
 XX
 DE Delta 1-46 urokinase.
 XX
 KW Human; des-epidermal growth factor homologous plasminogen activator; uPA;
 KW liver membrane; reduced affinity; EGF homologous; thrombosis;
 KW thrombolytic; increased half-life; urokinase.
 XX
 OS Homo sapiens; (engineered).
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "amino acids 1-46 of wild-type urokinase have been
 FT deleted"
 XX
 PN US5376547-A.
 XX
 XX 27-DEC-1994.
 XX
 XX 29-JAN-1988; 88US-00150267.
 XX
 XX 30-JAN-1987; 87US-00008795.
 XX
 XX (AMH) AMERICAN HOME PROD CORP.
 XX
 XX Hung PP, Lee SL, Kalyan NK;
 XX WPI; 1995-043464/06.
 XX
 XX New modified plasminogen activator cpds. - having regions removed to
 PT reduce affinity for liver membranes and increase circulation half-life.
 XX
 PS Claim 1; Page 2; 26pp; English.
 XX
 CC Amino acid residues 1-46 contain the EGF region of human urokinase.
 CC Deletion of this region results in a plasminogen activator with reduced
 CC affinity for liver cell membranes; the mutant protein is not cleared from
 CC the circulation as rapidly as is wild-type tPA. The specification only
 CC gives the sequence around the deletion and not the full-length sequence
 CC of "delta 1-46 urokinase"; the sequence in AAR68854 has been obtained by
 CC amending a previously disclosed wild-type human urokinase sequence (from
 CC W09501427) according to the description given in Example 3. (Updated on
 CC 25-MAR-2003 to correct PF field.) (Updated on 16-OCT-2003 to standardise
 CC OS field)
 XX
 SQ Sequence 365 AA;
 Query Match 100.0%; Score 554; DB 2; Length 365;
 Best Local Similarity 100.0%; Pred. No. 1.8e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNCRPN 60
 Db 2 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNCRPN 61
 Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPEE 96
 Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPEE 97
 RESULT 6
 AA013635
 ID AA013635 standard; protein; 378 AA.
 XX
 AC AA013635;
 XX
 DT 04-JUN-1997 (first entry)
 XX
 DE Human prourokinase variant lacking entire EGF domain.
 XX
 KW Human; prourokinase; hPUK; variant; half-life; increase; EGF;
 KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.

XX Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 1..9
 FT /note= "residues 1-9 of native hPUK"
 FT Region 10..378
 FT /note= "residues 43-411 of native hPUK"
 FT Misc-difference 144
 FT /note= "corresponds to TAC codon"
 XX
 PN EP398361-A.
 XX
 PD 22-NOV-1990.
 XX
 PF 18-MAY-1990; 90EP-00109472.
 XX
 PR 18-MAY-1989; 89JP-00126433.
 PR 22-FEB-1990; 90JP-00042020.
 XX
 PA (GREC) GREEN CROSS CORP.
 XX
 PI Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;
 PI Airmura H;
 XX
 DR WPI; 1990-350146/47.
 DR N-PSDB; AAT61672.
 XX
 XX Human pro-urokinase variants - deficient in loop regions of epidermal
 PT growth factor, showing long blood half-life, as fibrinolytic agent.
 XX
 PS Claim 1; Page; 22pp; English.
 XX
 CC New variants of human prourokinase (hPUK) comprise a hPUK deficient in
 CC (i) at least part of the first loop region of the epidermal growth factor
 CC (EGF) domain; (ii) at least part of the first loop and at least part of
 CC the second loop; or (iii) at least part of the third loop. The hPUK
 CC variants show an increased blood half-life comparable to that of the
 CC whole EGF domain-deficient hPUK variant and urokinase while retaining the
 CC same properties as those of hPUK. They have potent thrombolytic activity
 CC and very little tendency to cause spontaneous bleeding. The present
 CC sequence represents a specific variant of hPUK which lacks the entire EGF
 CC domain; the sequence does not appear in the specification and has been
 CC created using the wild-type hPUK sequence and the junction sequence after
 CC deletion, both of which are given (in Fig 1 and in Fig 2(3)),
 CC respectively)
 XX
 SQ Sequence 378 AA;
 Query Match 100.0%; Score 554; DB 2; Length 378;
 Best Local Similarity 100.0%; Pred. No. 1.9e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNCRPN 60
 Db 15 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNCRPN 74
 Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPEE 96
 Db 75 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPEE 110
 RESULT 7
 AAR66266
 ID AAR66266 standard; protein; 386 AA.
 XX
 AC AAR66266;
 XX
 DT 25-MAR-2003 (revised)
 DT 17-AUG-1995 (first entry)
 XX
 DE Bifunctional urokinase variant M33.

```

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Region 1..365
XX /label= M4
XX /note= "unglycosylated prourokinase(Ser47-Leu411)"
XX
XX Disulfide-bond 4..85
XX Disulfide-bond 25..67
XX Disulfide-bond 56..80
XX Disulfide-bond 102..233
XX Disulfide-bond 143..159
XX Disulfide-bond 151..222
XX Disulfide-bond 247..316
XX Disulfide-bond 279..295
XX Disulfide-bond 306..334
XX Disulfide-bond 365..366
XX Region /label= X1
XX /notes "peptide bond"
XX Region 366..386
XX /label= Y1
XX
XX DE4323754-Cl.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-04323754.
XX
XX 15-JUL-1993; 93DE-04323754.
XX
XX (CHEF ) GRUENTHAL GMBH.
XX
XX Steffens GJ; Wnendt S, Schneider J, Heinzl-Wieland R;
XX Saunders DJ;
XX
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with improved
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and
XX cerebral infarct, pulmonary embolism, etc.
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
XX are claimed (see features table). Sequences AAR6244-R6266 are specific
XX examples of such derivs. which have both improved fibrinolytic and
XX thrombin-inhibiting activities, compared to known plasminogen activators
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XX to correct PN field.)
XX
XX Sequence 386 AA;
XX
Query Match 100.0%; Score 554; DB 2; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.9e-44; Indels 0; Gaps 0;
Matches 96; Conservative 0; Mismatches 0;
QY 1 KTCYEGNGHYRKGASTDTMGRCPLPWSATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHYRKGASTDTMGRCPLPWSATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKSPSPPEE 96
Db 62 RRRPWCYVQGLKPLVQECMVHDCADGKSPSPPEE 97
XX
XX AAW13636 standard; protein; 389 AA.
XX
RESULT 8
AAW13636
ID AAW13636

```

```

XX AAW13636;
XX
XX 04-JUN-1997 (first entry)
XX
XX Human prourokinase variant lacking EGF domain loops 1 and 2.
XX
XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;
XX epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Region 1..10
XX /note= "residues 1-10 of native hPUK"
XX
XX Region 11..389
XX /note= "residues 33-411 of native hPUK"
XX
XX Misc-difference 155
XX /note= "corresponds to TAC codon"
XX
XX EP398361-A.
XX
XX 22-NOV-1990.
XX
XX 18-MAY-1990; 90EP-00109472.
XX
XX 18-MAY-1989; 89JP-00126433.
XX
XX 22-FEB-1990; 90JP-00042020.
XX
XX (GREC ) GREEN CROSS CORP.
XX
XX Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;
XX Airmura H;
XX
XX WPI; 1990-350146/47.
XX
XX N-PSDB; AAT61673.
XX
XX Human pro-urokinase variants - deficient in loop regions of epidermal
XX growth factor, showing long blood half-life, as fibrinolytic agent.
XX
XX Claim 6; Page; 22pp; English.
XX
XX New variants of human prourokinase (hPUK) comprise a hPUK deficient in
XX (i) at least part of the first loop region of the epidermal growth factor
XX (EGF) domain; (ii) at least part of the first loop and at least part of
XX the second loop; or (iii) at least part of the third loop. The hPUK
XX variants show an increased blood half-life comparable to that of the
XX whole EGF domain-deficient hPUK variant and urokinase while retaining the
XX same properties as those of hPUK. They have potent thrombolytic activity
XX and very little tendency to cause spontaneous bleeding. The present
XX sequence represents a specific variant of hPUK which lacks loops 1 and 2
XX of the EGF domain; the sequence does not appear in the specification and
XX has been created using the wild-type hPUK sequence and the junction
XX sequence after deletion, both of which are given (in Fig 1 and on page 8,
XX respectively)
XX
XX Sequence 389 AA;
XX
Query Match 100.0%; Score 554; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 1.9e-44; Indels 0; Gaps 0;
Matches 96; Conservative 0; Mismatches 0;
QY 1 KTCYEGNGHYRKGASTDTMGRCPLPWSATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 60
Db 26 KTCYEGNGHYRKGASTDTMGRCPLPWSATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 85
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKSPSPPEE 96
Db 86 RRRPWCYVQGLKPLVQECMVHDCADGKSPSPPEE 121
XX
XX RESULT 9

```


AAR66245
 ID AAR66245 standard; protein; 390 AA.
 XX
 AC AAR66245;
 XX
 DT 25-MAR-2003 (revised)
 DT 17-AUG-1995 (first entry)
 XX
 DE Bifunctional urokinase variant M12.
 XX
 KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX
 OS Synthetic.
 XX
 PH Key
 FT Location/Qualifiers
 FT 1. 365
 FT /label= M4
 FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
 FT Disulfide-bond 4. .85
 FT Disulfide-bond 25. .67
 FT Disulfide-bond 56. .80
 FT Disulfide-bond 102. .233
 FT Disulfide-bond 143. .159
 FT Disulfide-bond 151. .222
 FT Disulfide-bond 247. .316
 FT Disulfide-bond 279. .295
 FT Disulfide-bond 306. .334
 FT Disulfide-bond 366. .371
 FT Region /label= X1
 FT Region 372. .390
 FT /label= Y1
 XX
 PN DE4323754-C1.
 XX
 PD 01-DEC-1994.
 XX
 PF 15-JUL-1993; 93DE-04323754.
 XX
 PR 15-JUL-1993; 93DE-04323754.
 XX
 PA (CHEF) GRUENTHAL GMBH.
 XX
 XX Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;
 PI Saunders DJ;
 XX
 DR WPI; 1995-015191/03.
 XX
 PT New bifunctional urokinase derivs and related plasmids - with improved
 PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and
 PT cerebral infarct, pulmonary embolism, etc.
 XX
 PS Example 1; Page 10 and Fig 1; 34pp; German.
 XX
 CC Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
 CC are claimed (see features table). Sequences AAR66244-R66266 are specific
 CC examples of such derivs. which have both improved fibrinolytic and
 CC thrombin-inhibiting activities, compared to known plasminogen activators
 CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
 CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
 CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
 CC to correct PN field.)
 XX
 SQ Sequence 390 AA;

Query Match
 Best Local Similarity 100.0%; Score 554; DB 2; Length 390;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHYRGKASTDTWGRPCLPWNSATVLCQTYHAHRSDALQGLGKKNYCRNPDN 60
 DB 2 KTCYEGNGHYRGKASTDTWGRPCLPWNSATVLCQTYHAHRSDALQGLGKKNYCRNPDN 61

QY 61 RRRPWCYVQGLKPLVQECMWHDCADGKKFSSPPEE 96
 DB 62 RRRPWCYVQGLKPLVQECMWHDCADGKKFSSPPEE 97
 RESULT 10
 AAR66247
 ID AAR66247 standard; protein; 390 AA.
 XX
 AC AAR66247;
 XX
 DT 25-MAR-2003 (revised)
 DT 17-AUG-1995 (first entry)
 XX
 DE Bifunctional urokinase variant M14.
 XX
 KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX
 OS Synthetic.
 XX
 PH Key
 FT Location/Qualifiers
 FT 1. 365
 FT /label= M4
 FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
 FT Disulfide-bond 4. .85
 FT Disulfide-bond 25. .67
 FT Disulfide-bond 56. .80
 FT Disulfide-bond 102. .233
 FT Disulfide-bond 143. .159
 FT Disulfide-bond 151. .222
 FT Disulfide-bond 247. .316
 FT Disulfide-bond 279. .295
 FT Disulfide-bond 306. .334
 FT Disulfide-bond 366. .371
 FT Region /label= X1
 FT Region 372. .390
 FT /label= Y1
 XX
 PN DE4323754-C1.
 XX
 PD 01-DEC-1994.
 XX
 PF 15-JUL-1993; 93DE-04323754.
 XX
 PR 15-JUL-1993; 93DE-04323754.
 XX
 PA (CHEF) GRUENTHAL GMBH.
 XX
 XX Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;
 PI Saunders DJ;
 XX
 DR WPI; 1995-015191/03.
 XX
 PT New bifunctional urokinase derivs and related plasmids - with improved
 PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and
 PT cerebral infarct, pulmonary embolism, etc.
 XX
 PS Example 1; Page 10 and Fig 1; 34pp; German.
 XX
 CC Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
 CC are claimed (see features table). Sequences AAR66244-R66266 are specific
 CC examples of such derivs. which have both improved fibrinolytic and
 CC thrombin-inhibiting activities, compared to known plasminogen activators
 CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
 CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
 CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
 CC to correct PN field.)
 XX
 SQ Sequence 390 AA;

Query Match
 Best Local Similarity 100.0%; Score 554; DB 2; Length 390;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDGTGRCPLPWSATVLOQTYHAHRSALQGLGKKNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDGTGRCPLPWSATVLOQTYHAHRSALQGLGKKNYCRNPDN 61

QY 61 RRRPCWYVQVGLKPLVQECWVHDCADGKKPSPPEE 96
Db 62 RRRPCWYVQVGLKPLVQECWVHDCADGKKPSPPEE 97

RESULT 11
AAR66260
ID AAR66260 standard; protein; 392 AA.
XX AAR66260;
XX 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M27.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX Synthetic.
XX Key Location/Qualifiers
FH Region 1..365 /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region /label= Y1
XX DE4323754-Cl.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-04323754.
XX 15-JUL-1993; 93DE-04323754.
XX (CHEF) GRUENTHAL GMBH.
XX Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;
PI Saunders DJ;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with improved
FT fibrinolytic and thrombin inhibiting activities, for treating cardiac and
PT cerebral infarct, pulmonary embolism, etc.
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
CC are claimed (see features table). Sequences AAR66264-R66266 are specific
CC examples of such derivs. which have both improved fibrinolytic and
CC thrombin-inhibiting activities, compared to known plasminogen activators
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003

CC to correct PN field.)
XX
SQ Sequence 392 AA;
Query Match 100.0%; Score 554; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.9e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDGTGRCPLPWSATVLOQTYHAHRSALQGLGKKNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDGTGRCPLPWSATVLOQTYHAHRSALQGLGKKNYCRNPDN 61

QY 61 RRRPCWYVQVGLKPLVQECWVHDCADGKKPSPPEE 96
Db 62 RRRPCWYVQVGLKPLVQECWVHDCADGKKPSPPEE 97

RESULT 12
AAR66264
ID AAR66264 standard; protein; 392 AA.
XX AAR66264;
XX 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M31.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX Synthetic.
XX Key Location/Qualifiers
FH Region 1..365 /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region /label= Y1
XX DE4323754-Cl.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-04323754.
XX 15-JUL-1993; 93DE-04323754.
XX (CHEF) GRUENTHAL GMBH.
XX Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;
PI Saunders DJ;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with improved
FT fibrinolytic and thrombin inhibiting activities, for treating cardiac and
PT cerebral infarct, pulmonary embolism, etc.
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
CC are claimed (see features table). Sequences AAR66264-R66266 are specific
CC examples of such derivs. which have both improved fibrinolytic and
CC thrombin-inhibiting activities, compared to known plasminogen activators
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003

CC are claimed (see features table). Sequences AAR66244-R66266 are specific
CC examples of such derivs. which have both improved fibrinolytic and
CC thrombin-inhibiting activities, compared to known plasminogen activators
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
CC to correct PN field.)
XX
SQ Sequence 392 AA;

Query Match 100.0%; Score 554; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.9e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQQTVAHRSALQLGLGKKNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQQTVAHRSALQLGLGKKNYCRNPDN 61
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKPKSPPEE 96
DB 62 RRRPWCYVQGLKPLVQECMVHDCADGKPKSPPEE 97

RESULT 13

AAR66255
ID AAR66255 standard; protein; 392 AA.

XX AAR66255;

DT 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)

DE Bifunctional urokinase variant M22.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.

OS Synthetic.

FH Key Location/Qualifiers

FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..371
FT /label= X1
FT Region 372..392
FT /label= Y1

XX DB4323754-C1.

XX 01-DEC-1994.

XX 15-JUL-1993; 93DE-04323754.

XX 15-JUL-1993; 93DE-04323754.

XX (CHEF) GRUENTHAL GMBH.

XX Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;

PI Saunders DJ;

XX WPI; 1995-015191/03.

XX New bifunctional urokinase derivs and related plasmids - with improved

PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and
PT cerebral infarct, pulmonary embolism, etc.

PS Example 1; Page 10 and Fig 1; 34pp; German.

XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
CC are claimed (see features table). Sequences AAR66244-R66266 are specific
CC examples of such derivs. which have both improved fibrinolytic and
CC thrombin-inhibiting activities, compared to known plasminogen activators
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
CC to correct PN field.)
XX

SQ Sequence 392 AA;

Query Match 100.0%; Score 554; DB 2; Length 392;

Best Local Similarity 100.0%; Pred. No. 1.9e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQQTVAHRSALQLGLGKKNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQQTVAHRSALQLGLGKKNYCRNPDN 61
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKPKSPPEE 96
DB 62 RRRPWCYVQGLKPLVQECMVHDCADGKPKSPPEE 97

RESULT 14

AAR66259
ID AAR66259 standard; protein; 392 AA.

XX AAR66259;

DT 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)

DE Bifunctional urokinase variant M26.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.

OS Synthetic.

FH Key Location/Qualifiers

FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..371
FT /label= X1
FT Region 372..392
FT /label= Y1

XX DE4323754-C1.

XX 01-DEC-1994.

XX 15-JUL-1993; 93DE-04323754.

XX 15-JUL-1993; 93DE-04323754.

XX (CHEF) GRUENTHAL GMBH.

PI Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;
 PI Saunders DJ;
 XX WPI; 1995-015191/03.
 XX New bifunctional urokinase derivs and related plasmids - with improved
 PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and
 PT cerebral infarct, pulmonary embolism, etc.
 XX Example 1; Page 11 and Fig 1; 34pp; German.
 XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
 CC are claimed (see features table). Sequences AAR6244-R6266 are specific
 CC examples of such derivs. which have both improved fibrinolytic and
 CC thrombin-inhibiting activities, compared to known plasminogen activators
 CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
 CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
 CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
 CC to correct PN field.)
 XX Sequence 392 AA;
 SQ Query Match 100.0%; Score 554; DB 2; Length 392;
 Best Local Similarity 100.0%; Pred. No. 1.9e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
 DB 2 KTCYEGNGHFGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
 DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 97
 RESULT 15
 AAR6261
 ID AAR6261 standard; protein; 392 AA.
 XX AAR6261;
 XX 25-MAR-2003 (revised)
 DT 17-AUG-1995 (first entry)
 XX Bifunctional urokinase variant M28.
 DE fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX Synthetic.
 OS Key Location/Qualifiers
 FH Region 1..365
 FT /label= M4
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Region 366..371
 FT /label= X1
 FT Region 372..392
 XX DB4323754-C1.
 XX 01-DEC-1994.
 XX

PF 15-JUL-1993; 93DE-04323754.
 XX 15-JUL-1993; 93DE-04323754.
 XX (CHEF) GRUENENTHAL GMBH.
 XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;
 PI Saunders DJ;
 XX WPI; 1995-015191/03.
 XX New bifunctional urokinase derivs and related plasmids - with improved
 PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and
 PT cerebral infarct, pulmonary embolism, etc.
 XX Example 1; Page 11 and Fig 1; 34pp; German.
 XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
 CC are claimed (see features table). Sequences AAR6244-R6266 are specific
 CC examples of such derivs. which have both improved fibrinolytic and
 CC thrombin-inhibiting activities, compared to known plasminogen activators
 CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
 CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
 CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
 CC to correct PN field.)
 XX Sequence 392 AA;
 SQ Query Match 100.0%; Score 554; DB 2; Length 392;
 Best Local Similarity 100.0%; Pred. No. 1.9e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
 DB 2 KTCYEGNGHFGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
 DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 97
 RESULT 16
 AAR6258
 ID AAR6258 standard; protein; 392 AA.
 XX AAR6258;
 XX 25-MAR-2003 (revised)
 DT 17-AUG-1995 (first entry)
 XX Bifunctional urokinase variant M25.
 DE fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX Synthetic.
 OS Key Location/Qualifiers
 FH Region 1..365
 FT /label= M4
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Region 366..371
 FT /label= X1
 FT Region 372..392
 XX

```

FT XX /label= Y1
PN DE4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-04323754.
XX
XX 15-JUL-1993; 93DE-04323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;
XX Saunders DJ;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with improved
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and
XX cerebral infarct, pulmonary embolism, etc.
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
XX are claimed (see features table). Sequences AAR6244-R66266 are specific
XX examples of such derivs. which have both improved fibrinolytic and
XX thrombin-inhibiting activities, compared to known plasminogen activators
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XX to correct PN field.)
XX
XX Sequence 392 AA;
XX
XX Query Match 100.0%; Score 554; DB 2; Length 392;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-44;
XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 KTCYEGNGHFGKASDTMGRCPLPWSATVLQQTYYHAHRSDALQLGLGKHNCRNPDN 60
XX 2 KTCYEGNGHFGKASDTMGRCPLPWSATVLQQTYYHAHRSDALQLGLGKHNCRNPDN 61
XX
XX Db 61 RRRPWCYVQGLKPLVQECMVHDCADGKFPSPPEE 96
XX 62 RRRPWCYVQGLKPLVQECMVHDCADGKFPSPPEE 97
XX
XX RESULT 17
XX AAR66263
XX ID AAR66263 standard; protein; 392 AA.
XX
XX AC AAR66263;
XX
XX DT 25-MAR-2003 (revised)
XX DT 17-AUG-1995 (first entry)
XX
XX DE Bifunctional urokinase variant M30.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Region 1. .365
XX /label= M4
XX /note= "unglycosylated prourokinase(Ser47-Leu411)"
XX
XX Disulfide-bond 4. .85
XX Disulfide-bond 25. .67
XX Disulfide-bond 56. .80
XX Disulfide-bond 102. .233
XX Disulfide-bond 143. .159
XX Disulfide-bond 151. .222

```

```

FT XX Disulfide-bond 247. .316
FT XX Disulfide-bond 279. .295
FT XX Disulfide-bond 306. .334
FT XX Region 366. .371
FT XX /label= X1
FT XX /label= Y1
XX
XX DE4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-04323754.
XX
XX 15-JUL-1993; 93DE-04323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;
XX Saunders DJ;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with improved
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and
XX cerebral infarct, pulmonary embolism, etc.
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
XX are claimed (see features table). Sequences AAR6244-R66266 are specific
XX examples of such derivs. which have both improved fibrinolytic and
XX thrombin-inhibiting activities, compared to known plasminogen activators
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XX to correct PN field.)
XX
XX Sequence 392 AA;
XX
XX Query Match 100.0%; Score 554; DB 2; Length 392;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-44;
XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 KTCYEGNGHFGKASDTMGRCPLPWSATVLQQTYYHAHRSDALQLGLGKHNCRNPDN 60
XX 2 KTCYEGNGHFGKASDTMGRCPLPWSATVLQQTYYHAHRSDALQLGLGKHNCRNPDN 61
XX
XX Db 61 RRRPWCYVQGLKPLVQECMVHDCADGKFPSPPEE 96
XX 62 RRRPWCYVQGLKPLVQECMVHDCADGKFPSPPEE 97
XX
XX RESULT 18
XX AAR66254
XX ID AAR66254 standard; protein; 392 AA.
XX
XX AC AAR66254;
XX
XX DT 25-MAR-2003 (revised)
XX DT 17-AUG-1995 (first entry)
XX
XX DE Bifunctional urokinase variant M21.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Region 1. .365
XX /label= M4
XX /note= "unglycosylated prourokinase(Ser47-Leu411)"
XX
XX Disulfide-bond 247. .316
XX Disulfide-bond 279. .295
XX Disulfide-bond 306. .334
XX Region 366. .371
XX /label= X1
XX /label= Y1

```

```

FT Disulfide-bond 4. .85
FT Disulfide-bond 25. .67
FT Disulfide-bond 56. .80
FT Disulfide-bond 102. .233
FT Disulfide-bond 143. .159
FT Disulfide-bond 151. .222
FT Disulfide-bond 247. .316
FT Disulfide-bond 279. .295
FT Disulfide-bond 306. .334
FT Disulfide-bond 366. .371
FT Region
FT /label= X1
FT Region
FT /label= Y1
XX DE4323754-Cl.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-04323754.
XX 15-JUL-1993; 93DE-04323754.
XX (CHEF ) GRUENENTHAL GMBH.
XX Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;
XX Saunders DU;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with improved
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and
XX cerebral infarct, pulmonary embolism, etc.
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
XX are claimed (see features table). Sequences AAR66244-R66266 are specific
XX examples of such derivs. which have both improved fibrinolytic and
XX thrombin-inhibiting activities, compared to known plasminogen activators
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XX to correct PN field.)
XX Sequence 392 AA;
XX Query Match 100.0%; Score 554; DB 2; Length 392;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-44;
XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 97
RESULT 19
AAR66256
ID AAR66256 standard; protein; 392 AA.
XX AAR66256;
XX 25-MAR-2003 (revised)
XX 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M23.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.

```

```

OS Synthetic.
XX Key
XX Region
XX Location/Qualifiers
XX 1. .365
XX /label= M4
XX /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4. .85
FT Disulfide-bond 25. .67
FT Disulfide-bond 56. .80
FT Disulfide-bond 102. .233
FT Disulfide-bond 143. .159
FT Disulfide-bond 151. .222
FT Disulfide-bond 247. .316
FT Disulfide-bond 279. .295
FT Disulfide-bond 306. .334
FT Disulfide-bond 366. .371
FT Region
FT /label= X1
FT Region
FT /label= Y1
XX DE4323754-Cl.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-04323754.
XX 15-JUL-1993; 93DE-04323754.
XX (CHEF ) GRUENENTHAL GMBH.
XX Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;
XX Saunders DU;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with improved
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and
XX cerebral infarct, pulmonary embolism, etc.
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
XX are claimed (see features table). Sequences AAR66244-R66266 are specific
XX examples of such derivs. which have both improved fibrinolytic and
XX thrombin-inhibiting activities, compared to known plasminogen activators
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XX to correct PN field.)
XX Sequence 392 AA;
XX Query Match 100.0%; Score 554; DB 2; Length 392;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-44;
XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 97
RESULT 20
AAR66257
ID AAR66257 standard; protein; 392 AA.
XX AAR66257;
XX 25-MAR-2003 (revised)
XX 17-AUG-1995 (first entry)

```

XX Bifunctional urokinase variant M24.
DE
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
XX
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region /label= Y1
XX
XX DE4323754-Cl.
XX
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-04323754.
XX
XX 15-JUL-1993; 93DE-04323754.
XX (CHEF) GRUENENTHAL GMBH.
XX
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;
PI Saunders DJ;
XX
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with improved
PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and
PT cerebral infarct, pulmonary embolism, etc.
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
CC are claimed (see features table). Sequences AAR66244-R66266 are specific
CC examples of such derivs. which have both improved fibrinolytic and
CC thrombin-inhibiting activities, compared to known plasminogen activators
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XX to correct PN field.)
XX
SQ Sequence 392 AA;
Query Match 100.0%; Score 554; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.9e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWSATVLCQTYHAHRSALQGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHYRGKASTDTMGRPCLPWSATVLCQTYHAHRSALQGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 97

RESULT 21
AAR66251

ID AAR66251 standard; protein; 393 AA.
XX
XX AAR66251;
XX
DT 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M18.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..372
FT Region /label= X1
FT Region /label= Y1
XX
XX DE4323754-Cl.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-04323754.
XX
XX 15-JUL-1993; 93DE-04323754.
XX (CHEF) GRUENENTHAL GMBH.
XX
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;
PI Saunders DJ;
XX
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with improved
PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and
PT cerebral infarct, pulmonary embolism, etc.
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
CC are claimed (see features table). Sequences AAR66244-R66266 are specific
CC examples of such derivs. which have both improved fibrinolytic and
CC thrombin-inhibiting activities, compared to known plasminogen activators
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XX to correct PN field.)
XX
SQ Sequence 393 AA;
Query Match 100.0%; Score 554; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.9e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWSATVLCQTYHAHRSALQGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHYRGKASTDTMGRPCLPWSATVLCQTYHAHRSALQGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 97

Db 62 RRRPWCYVQVGLKPLVQECWVHDCADGKKPSPPEE 97

RESULT 22

AAR66253
ID AAR66253 standard; protein; 393 AA.

XX AAR66253;
AC AAR66253;

DT 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)

XX XX

DE Bifunctional urokinase variant M20.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.

XX Synthetic.

XX Key Location/Qualifiers
FH 1. 365
FT Region /label= M4

FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

FT Disulfide-bond 4. .85

FT Disulfide-bond 25. .67

FT Disulfide-bond 56. .80

FT Disulfide-bond 102. .233

FT Disulfide-bond 143. .159

FT Disulfide-bond 151. .222

FT Disulfide-bond 247. .316

FT Disulfide-bond 279. .295

FT Disulfide-bond 306. .334

FT Disulfide-bond 366. .372

FT Region /label= X1

FT Region 373. .393

FT /label= Y1

XX DE4323754-C1

XX PD 01-DEC-1994.

XX 15-JUL-1993; 93DE-04323754.

XX 15-JUL-1993; 93DE-04323754.

XX (CHEF) GRUENENTHAL GMBH.

XX Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;

XX Saunders DJ;

XX WPI; 1995-015191/03.

XX New bifunctional urokinase derivs and related plasmids - with improved

XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and

XX cerebral infarct, pulmonary embolism, etc.

XX Example 1; Page 10 and Fig 1; 34pp; German.

XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1

XX are claimed (see features table). Sequences AAR66244-R66266 are specific

XX examples of such derivs. which have both improved fibrinolytic and

XX thrombin-inhibiting activities, compared to known plasminogen activators

XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,

XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac

XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003

XX to correct PN field.)

XX SQ Sequence 393 AA;

Query Match 100.0%; Score 554; DB 2; Length 393;

Best Local Similarity 100.0%; Pred. No. 1.9e-44;

Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWNSATVLQOTYHAHRSDALQGLGKGYCRNPON 60
Db 2 KTCYEGNGHFYRGKASTDTMGPRCLPWNSATVLQOTYHAHRSDALQGLGKGYCRNPON 61

QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGKKPSPPEE 96
Db 62 RRRPWCYVQVGLKPLVQECWVHDCADGKKPSPPEE 97

RESULT 23

AAR66249

ID AAR66249 standard; protein; 393 AA.

XX AAR66249;

DT 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)

XX XX

DE Bifunctional urokinase variant M16.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.

XX Synthetic.

XX Key Location/Qualifiers
FH 1. 365
FT Region /label= M4

FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

FT Disulfide-bond 4. .85

FT Disulfide-bond 25. .67

FT Disulfide-bond 56. .80

FT Disulfide-bond 102. .233

FT Disulfide-bond 143. .159

FT Disulfide-bond 151. .222

FT Disulfide-bond 247. .316

FT Disulfide-bond 279. .295

FT Disulfide-bond 306. .334

FT Disulfide-bond 366. .372

FT Region /label= X1

FT Region 373. .393

FT /label= Y1

XX DE4323754-C1.

XX PD 01-DEC-1994.

XX 15-JUL-1993; 93DE-04323754.

XX 15-JUL-1993; 93DE-04323754.

XX (CHEF) GRUENENTHAL GMBH.

XX Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;

XX Saunders DJ;

XX WPI; 1995-015191/03.

XX New bifunctional urokinase derivs and related plasmids - with improved

XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and

XX cerebral infarct, pulmonary embolism, etc.

XX Example 1; Page 10 and Fig 1; 34pp; German.

XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1

XX are claimed (see features table). Sequences AAR66244-R66266 are specific

XX examples of such derivs. which have both improved fibrinolytic and

XX thrombin-inhibiting activities, compared to known plasminogen activators

XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,

XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac

XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003

XX to correct PN field.)

XX Sequence 393 AA;
SQ Query Match 100.0%; Score 554; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.9e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPNSATVLQOYTHAHRSDALQLGLGKKNYCRNPDN 60
DB 2 KTCYEGNGHYRGKASTDTMGRCPLPNSATVLQOYTHAHRSDALQLGLGKKNYCRNPDN 61

QY 61 RRPWCYVQVGLKPLVQECWVHDCADGKPKSPPEE 96
DB 62 RRPWCYVQVGLKPLVQECWVHDCADGKPKSPPEE 97

RESULT 24
AAR66252
ID AAR66252 standard; protein; 393 AA.
XX AC AAR66252;
XX DT 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)
XX DB Bifunctional urokinase variant M19.
XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX OS Synthetic.

XX Key Location/Qualifiers
FH Region 1.365
FT /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4.85
FT Disulfide-bond 25.87
FT Disulfide-bond 56.80
FT Disulfide-bond 102.233
FT Disulfide-bond 143.159
FT Disulfide-bond 151.222
FT Disulfide-bond 247.316
FT Disulfide-bond 279.295
FT Disulfide-bond 306.334
FT Region 366.372
FT /label= X1
FT Region 373.393
FT /label= Y1

PN DE4323754-C1.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-04323754.
XX 15-JUL-1993; 93DE-04323754.
XX (CHEP) GRUENTHAL GMBH.
XX Steffens GJ; Wndt S, Schneider J, Heinzl-Wieland R;
XX Saunders DJ;
XX WPI; 1995-015191/03.
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
XX are claimed (see features table). Sequences AAR66244-R66286 are specific

CC examples of such derivs. which have both improved fibrinolytic and
CC thrombin-inhibiting activities, compared to known plasminogen activators
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
CC to correct PN field.)
XX Sequence 393 AA;
SQ Query Match 100.0%; Score 554; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.9e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPNSATVLQOYTHAHRSDALQLGLGKKNYCRNPDN 60
DB 2 KTCYEGNGHYRGKASTDTMGRCPLPNSATVLQOYTHAHRSDALQLGLGKKNYCRNPDN 61

QY 61 RRPWCYVQVGLKPLVQECWVHDCADGKPKSPPEE 96
DB 62 RRPWCYVQVGLKPLVQECWVHDCADGKPKSPPEE 97

RESULT 25
AAR66244
ID AAR66244 standard; protein; 393 AA.
XX AC AAR66244;
XX DT 25-MAR-2003 (revised)
DT 22-AUG-1995 (first entry)
XX DE Bifunctional urokinase variant M11.
XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX OS Synthetic.

XX Key Location/Qualifiers
FH Region 1.365
FT /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4.85
FT Disulfide-bond 25.87
FT Disulfide-bond 56.80
FT Disulfide-bond 102.233
FT Disulfide-bond 143.159
FT Disulfide-bond 151.222
FT Disulfide-bond 247.316
FT Disulfide-bond 279.295
FT Disulfide-bond 306.334
FT Region 366.372
FT /label= X1
FT Region 375.393
FT /label= Y1

PN DE4323754-C1.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-04323754.
XX 15-JUL-1993; 93DE-04323754.
XX (CHEP) GRUENTHAL GMBH.
XX Steffens GJ; Wndt S, Schneider J, Heinzl-Wieland R;
XX Saunders DJ;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with improved
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and
XX cerebral infarct, pulmonary embolism, etc.
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
XX are claimed (see features table). Sequences AAR66244-R66286 are specific

PT cerebral infarct, pulmonary embolism, etc.
XX
PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
CC Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
CC are claimed (see features table). Sequences AAR66244-R66266 are specific
CC examples of such derivs. which have both improved fibrinolytic and
CC thrombin-inhibiting activities, compared to known plasminogen activators
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
CC to correct PN field.)
XX
SQ Sequence 393 AA;
Query Match 100.0%; Score 554; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.9e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTNGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTNGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
QY 61 RRRPCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
DB 62 RRRPCYVQVGLKPLVQECMVHDCADGKKPSPPEE 97
RESULT 26
AAR66250
ID AAR66250 standard; protein; 393 AA.
XX
AC AAR66250;
XX
DT 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M17.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..372
FT Region /label= X1
FT Region 373..393
FT /label= Y1
XX
DE4323754-Cl.
XX
PD 01-DEC-1994.
XX
PF 15-JUL-1993; 93DE-04323754.
XX
PR 15-JUL-1993; 93DE-04323754.
XX (CHEF) GRUENENTHAL GMBH.
PA Steffens GJ, Wnerdt S, Schneider J, Heinzel-Wieland R;
PI

PI Saunders DU;
XX
DR WPI; 1995-015191/03.
XX
PT New bifunctional urokinase derivs and related plasmids - with improved
PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and
PT cerebral infarct, pulmonary embolism, etc.
XX
PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
CC Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
CC are claimed (see features table). Sequences AAR66244-R66266 are specific
CC examples of such derivs. which have both improved fibrinolytic and
CC thrombin-inhibiting activities, compared to known plasminogen activators
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
CC to correct PN field.)
XX
SQ Sequence 393 AA;
Query Match 100.0%; Score 554; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.9e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTNGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTNGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
QY 61 RRRPCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
DB 62 RRRPCYVQVGLKPLVQECMVHDCADGKKPSPPEE 97
RESULT 27
AAR99597
ID AAR99597 standard; protein; 393 AA.
XX
AC AAR99597;
XX
DT 05-DEC-1996 (first entry)
DT
DE Chimeric protein M38 encoded by pSE1.
XX
KW Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;
KW plasminogen activating sequence; fibrinolysis; infarction;
KW angina pectoris; deep vein thrombosis.
XX
OS Synthetic.
XX
PN EF714982-A2.
XX
PD 05-JUN-1996.
XX
PF 16-NOV-1995; 95EP-00118050.
XX
PR 30-NOV-1994; 94DE-04442665.
XX
PA (CHEF) GRUENENTHAL GMBH.
XX
PI Wnerdt S, Steffens GJ, Janocha E, Heinzel-Wieland R;
XX
DR WPI; 1996-269715/28.
XX
PT Chimeric protein contg. plasminogen activating sequence and thrombin-
PT inhibiting sequence - useful as thrombus-specific thrombolytic agent with
PT rapid action.
XX
PS Example 1; Page 21-22; 37pp; German.
XX
CC Example 1 describes the prodn. of plasmids pSE1 and pSE9 contg. a DNA
CC encoding a chimeric protein with fibrinolytic and thrombin-inhibiting
CC properties. pSE1 encodes the protein given in AAR99597 and pSE9 encodes

```

CC the protein given in AAR99596
XX
SQ Sequence 393 AA;

Query Match      100.0%; Score 554; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.9e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 60
Dd 3 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 62
QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 96
Dd 63 RRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 98

RESULT 28
AAR99596
ID AAR99596 standard; protein; 393 AA.
XX
AC AAR99596;
XX
DT 05-DEC-1996 (first entry)
XX
DE Chimeric protein M37 encoded by pSE9.
XX
KW Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;
KW plasminogen activating sequence; fibrinolysis; infarction;
KW angina pectoris; deep vein thrombosis.
XX
OS Synthetic.
XX
PN EP714982-A2.
XX
PD 05-JUN-1996.
XX
PF 16-NOV-1995; 95EP-00118050.
XX
PR 30-NOV-1994; 94DE-04442665.
XX
PA (CHEF ) GRUENENTHAL GMBH.
XX
PI Wnendt S, Steffens GJ, Janocha E, Heinzel-Wieland R;
XX WPI; 1996-269715/28.
XX
DR Chimeric protein contg. plasminogen activating sequence and thrombin-
PT inhibiting sequence - useful as thrombus-specific thrombolytic agent with
PT rapid action.
XX
PS Example 1; Page 19-20; 37pp; German.
XX
CC Example 1 describes the prodn. of plasmids pSE1 and pSE9 contg. a DNA
CC encoding a chimeric protein with fibrinolytic and thrombin-inhibiting
CC properties. pSE1 encodes the protein given in AAR99597 and pSE9 encodes
CC the protein given in AAR99596
XX
SQ Sequence 393 AA;

Query Match      100.0%; Score 554; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.9e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 60
Dd 3 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 62
QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 96
Dd 63 RRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 98

us-09-880-503-9.rag

```

RESULT 30
 AAR47902
 ID AAR47902 standard; protein; 395 AA.
 AC AAR47902;
 XX
 XX
 DT 13-JUL-1994 (first entry)
 DT Pro-urokinase derivative.
 DE
 DE Pro-urokinase; half-life; thrombolytic; thrombosis; fibrinolytic; factor.
 KW Homo sapiens.
 OS
 XX
 XX JP05336965-A.
 XX
 XX 21-DEC-1993.
 XX
 XX 17-OCT-1991; 91JP-00269615.
 XX
 XX 17-OCT-1991; 91JP-00269615.
 XX
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 XX WPI; 1994-030907/04.
 DR N-PSDB; AAQ55771.
 XX
 XX Novel human pro-urokinase derivs. having long half-life - with high
 FT thrombolytic activity; useful for treatment of thrombosis.
 PT
 XX
 XX Disclosure; Page 14; 29pp; Japanese.
 XX
 CC Sequences (AAQ55771-72) are pro-urokinase derivatives. The products have
 CC an inserted sugar moiety having an amino acid substituted, depleted or
 CC inserted variant around the thrombin cleavage site. They also have a long
 CC half-life allowing them to be used in the treatment of thrombosis
 XX
 XX Sequence 395 AA;
 SQ
 Query Match 100.0%; Score 554; DB 2; Length 395;
 Best Local Similarity 100.0%; Pred. No. 1.9e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 60
 Db 32 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 91
 QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGKPKSPSPPEE 96
 Db 92 RRRPWCYVQVGLKPLVQECWVHDCADGKPKSPSPPEE 127
 RESULT 31
 AAR66265
 ID AAR66265 standard; protein; 395 AA.
 AC AAR66265;
 XX
 XX 25-MAR-2003 (revised)
 DT 17-AUG-1995 (first entry)
 XX
 DE Bifunctional urokinase variant M32.
 KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT 1..365
 FT Region /label= M4
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

FT Disulfide-bond 4..85
 FT Disulfide-bond 25..57
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Disulfide-bond 366..371
 FT Region /label= X1
 FT Region 372..395
 FT /label= Y1
 XX
 XX DE4323754-C1.
 XX
 PD 01-DEC-1994.
 XX
 PD 15-JUL-1993; 93DE-04323754.
 XX
 PD 15-JUL-1993; 93DE-04323754.
 XX
 PA (CHEF) GRUENENTHAL GMBH.
 XX
 PI Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;
 PI Saunders DJ;
 XX
 DR WPI; 1995-015191/03.
 XX
 XX New bifunctional urokinase derivs and related plasmids - with improved
 FT fibrinolytic and thrombin inhibiting activities, for treating cardiac and
 PT cerebral infarct, pulmonary embolism, etc.
 XX
 XX Example 1; Page 11 and Fig 1; 34pp; German.
 XX
 CC Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
 CC are claimed (see features table). Sequences AAR66244-R66266 are specific
 CC examples of such derivs. which have both improved fibrinolytic and
 CC thrombin-inhibiting activities, compared to known plasminogen activators
 CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
 CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
 CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
 CC to correct PN field.)
 XX
 SQ Sequence 395 AA;
 Query Match 100.0%; Score 554; DB 2; Length 395;
 Best Local Similarity 100.0%; Pred. No. 1.9e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 60
 Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 61
 QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGKPKSPSPPEE 96
 Db 62 RRRPWCYVQVGLKPLVQECWVHDCADGKPKSPSPPEE 97
 RESULT 32
 AAR66262
 ID AAR66262 standard; protein; 395 AA.
 XX
 AC AAR66262;
 XX
 XX 25-MAR-2003 (revised)
 DT 17-AUG-1995 (first entry)
 XX
 DE Bifunctional urokinase variant M29.
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 KW
 XX

```

OS Synthetic.
XX Key Location/Qualifiers
XX Region 1. .365
XX /label= M4
XX /note= "unglycosylated prourokinase(Ser47-Leu411)"
XX Disulfide-bond 4. .85
XX Disulfide-bond 25. .67
XX Disulfide-bond 56. .80
XX Disulfide-bond 102. .233
XX Disulfide-bond 143. .159
XX Disulfide-bond 151. .222
XX Disulfide-bond 247. .316
XX Disulfide-bond 279. .295
XX Disulfide-bond 306. .334
XX Disulfide-bond 366. .371
XX Region /label= X1
XX Region /label= Y1
XX DE4323754-C1.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-04323754.
XX 15-JUL-1993; 93DE-04323754.
XX (CHEP ) GRUENENTHAL GMBH.
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;
XX Saunders DJ;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with improved
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and
XX cerebral infarct, pulmonary embolism, etc.
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
XX are claimed (see features table). Sequences AAR6244-R66266 are specific
XX examples of such derivs. which have both improved fibrinolytic and
XX thrombin-inhibiting activities, compared to known plasminogen activators
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XX to correct FN field.)
XX Sequence 395 AA;
XX Query Match 100.0%; Score 554; DB 2; Length 395;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-44;
XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 60
DB 2 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 97
RESULT 33
AAR66246
ID AAR66246 standard; protein; 396 AA.
XX
XX AAR66246;
XX
XX 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)

```

```

XX Bifunctional urokinase variant M13.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.
XX Synthetic.
XX Key Location/Qualifiers
XX Region 1. .365
XX /label= M4
XX /note= "unglycosylated prourokinase(Ser47-Leu411)"
XX Disulfide-bond 4. .85
XX Disulfide-bond 25. .67
XX Disulfide-bond 56. .80
XX Disulfide-bond 102. .233
XX Disulfide-bond 143. .159
XX Disulfide-bond 151. .222
XX Disulfide-bond 247. .316
XX Disulfide-bond 279. .295
XX Disulfide-bond 306. .334
XX Disulfide-bond 366. .377
XX Region /label= X1
XX Region /label= Y1
XX DE4323754-C1.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-04323754.
XX 15-JUL-1993; 93DE-04323754.
XX (CHEP ) GRUENENTHAL GMBH.
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;
XX Saunders DJ;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derive and related plasmids - with improved
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and
XX cerebral infarct, pulmonary embolism, etc.
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
XX are claimed (see features table). Sequences AAR6244-R66266 are specific
XX examples of such derivs. which have both improved fibrinolytic and
XX thrombin-inhibiting activities, compared to known plasminogen activators
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XX to correct FN field.)
XX Sequence 396 AA;
XX Query Match 100.0%; Score 554; DB 2; Length 396;
XX Best Local Similarity 100.0%; Pred. No. 2e-44;
XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 60
DB 2 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 97
RESULT 34
AAR66248

```

```

ID AAR66248 standard; protein; 397 AA.
XX
AC AAR66248;
XX
DT 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M15.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..378
FT Region /label= X1
FT Region /label= Y1
XX
PN DE4323754-CL.
XX
PD 01-DEC-1994.
XX
PF 15-JUL-1993. 93DE-04323754.
XX
PR 15-JUL-1993. 93DE-04323754.
XX
PA (CHEF ) GRUENENTHAL GMBH.
XX
XX
PI Steffens GU, Whendt S, Schneider J, Heinzl-Wieland R;
PI Saunders DJ;
XX
XX
DR WPI; 1995-015191/03.
XX
XX
PT New bifunctional urokinase derivs and related plasmids - with improved
PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and
PT cerebral infarct, pulmonary embolism, etc.
XX
PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX
CC Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
CC are claimed (see features table). Sequences AAR66244-R66266 are specific
CC examples of such derivs which have both improved fibrinolytic and
CC thrombin-inhibiting activities, compared to known plasminogen activators
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
CC to correct PN field.)
XX
SQ Sequence 397 AA;
Query Match 100.0%; Score 554; DB 2; Length 397;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHYCRNPDN 60
DB 2 KTCYEGNGHFRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHYCRNPDN 61
QY 61 RRRPCYVQVGLKPLVQECMWHDCADGKPKSPPEE 96

```

```

Db 62 RRRPCYVQVGLKPLVQECMWHDCADGKPKSPPEE 97
|||||
RESULT 35
AAW13637
ID AAW13637 standard; protein; 401 AA.
XX
AC AAW13637;
XX
DT 04-JUN-1997 (first entry)
XX
DE Human prourokinase variant lacking EGF domain loop 3.
XX
KW Human; prourokinase; hPUK; variant; half-life; increase; EGF;
KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Region 1..32
FT /note= "residues 1-32 of native hPUK"
FT Region 33..401
FT /note= "residues 43-411 of native hPUK"
FT Misc-difference 167
FT /note= "corresponds to TAC codon"
XX
PN EP398361-A.
XX
PD 22-NOV-1990.
XX
PF 18-MAY-1990; 90EP-00109472.
XX
PR 18-MAY-1989; 89JP-00126433.
PR 22-FEB-1990; 90JP-00042020.
XX
PA (GREC ) GREEN CROSS CORP.
XX
XX
PI Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;
PI Airmura H;
XX
DR WPI; 1990-350146/47.
DR N-PSDB; AAT61674.
XX
XX
PT Human pro-urokinase variants - deficient in loop regions of epidermal
PT growth factor, showing long blood half-life, as fibrinolytic agent.
XX
PS Claim 11; Page; 22pp; English.
XX
XX
CC New variants of human prourokinase (hPUK) comprise a hPUK deficient in
CC (i) at least part of the first loop region of the epidermal growth factor
CC (EGF) domain; (ii) at least part of the first loop and at least part of
CC the second loop; or (iii) at least part of the third loop. The hPUK
CC variants show an increased blood half-life comparable to that of the
CC whole EGF domain-deficient hPUK variant and urokinase while retaining the
CC same properties as those of hPUK. They have potent thrombolytic activity
CC and very little tendency to cause spontaneous bleeding. The present
CC sequence represents a specific variant of hPUK which lacks EGF domain
CC loop 3; the sequence does not appear in the specification and has been
CC created using the wild-type hPUK sequence and the junction sequence after
CC deletion, both of which are given (in Fig 1 and on page 8, respectively)
XX
SQ Sequence 401 AA;
Query Match 100.0%; Score 554; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHYCRNPDN 60
DB 38 KTCYEGNGHFRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHYCRNPDN 97

```

QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKKPSSPPEE 96
DB 98 RRRPWCYVQGLKPLVQECMVHDCADGKKPSSPPEE 133

RESULT 36
AAR23794
ID AAR23794 standard; protein; 410 AA.

XX AC AAR23794;
XX DT 03-NOV-1992 (first entry)
XX DE Prourokinase mutant F157D.
XX KW Substitution; animal; plasminogen activator; blood; clot.
XX OS Homo sapiens.

XX PN JP04091792-A.
XX PD 25-MAR-1992.
XX PF 07-AUG-1990; 90JP-00207659.
XX PR 07-AUG-1990; 90JP-00207659.

XX PA (TOYU) TOSCH CORP.
XX DR WPI; 1992-154820/19.
XX PT Vector contg. pro-urokinase encoding gene - includes mouse-IgG H chain E-mu enhancer, enhancer contg. SV40, early phase promoter and SV40 poly:adenylation site.

XX PS Claim 3; Page 1; 15pp; Japanese.
XX CC The sequence is that of mature prourokinase having the substitution mutation F157D. The sequence per se is not given in the specification, so the known sequence of prourokinase was used to demonstrate the mutant protein. The gene encoding prourokinase has been over-expressed by recombinant DNA technology and provides a method for the rapid, low cost prodn. of prourokinase, which is a low mol. wt. plasminogen activator used for dissolving blood clots. See also AAR23795

XX SQ Sequence 410 AA;

Query Match 100.0%; Score 554; DB 2; Length 410;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKKNYCRNPDN 60
DB 47 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKKNYCRNPDN 106

QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKKPSSPPEE 96
DB 107 RRRPWCYVQGLKPLVQECMVHDCADGKKPSSPPEE 142

RESULT 37
AAP50871
ID AAP50871 standard; protein; 411 AA.

XX AC AAP50871;
XX DT 30-NOV-1991 (first entry)

XX DE Sequence encoded by cDNA sequence for human urokinase zymogen (Japanese Patent Application No.37119/84).
XX KW Thrombolytic agent; plasminogen activator activity; fibrin affinity; enzyme.

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Cleavage-site 158..159
FT /note= "potential cleavage site which generates the two-chain form from the zymogen"

FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380

XX EPI39447-A.

XX PD 02-MAY-1985.

XX PF 07-SEP-1984; 84EP-00306117.

XX PR 13-SEP-1983; 83JP-00170354.

XX PR 17-OCT-1983; 83JP-00195051.

XX PA (GREG) GREEN CROSS CORP.

XX PI Kasai S, Arimura H, Mori K, Suyama T;

XX DR WPI; 1985-106530/18.

XX PT New urokinase zymogen - useful as thrombolytic agent.

XX PS Disclosure; Page 12; 30pp; English.

XX CC Zymogen AAP50871 is the inactive precursor form of human urokinase. Urokinase zymogen is cleaved into the two-chain form composed of characteristic urokinase H (molecular wt. of 30,000) and L (molecular wt. of 20,000) chains when treated with catalytic amounts of plasmin. The patentors claim a new urokinase zymogen which has mol. wt. ca. 50,000, a single chain molecular structure, and selective affinity for fibrin. It is a thrombolytic agent which manifests its plasminogen activator activity on cleavage by proteolytic enzymes (e.g. plasmin) and has higher affinity for fibrin than known forms of urokinase

XX SQ Sequence 411 AA;

Query Match 100.0%; Score 554; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKKNYCRNPDN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKKNYCRNPDN 107

QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKKPSSPPEE 96
DB 108 RRRPWCYVQGLKPLVQECMVHDCADGKKPSSPPEE 143

RESULT 38
AAR06244
ID AAR06244 standard; protein; 411 AA.

XX AC AAR06244;
XX DT 07-DEC-1990 (first entry)

XX DE Urokinase precursor protein.

XX KW Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis; myocardial infarction.

```
XX OS Homo sapiens.
XX PN EP380334-A.
XX PD 01-AUG-1990.
XX PF 25-JAN-1990; 90EP-00300772.
XX PR 27-JAN-1989; 89JP-00016406.
XX PA 17-MAY-1989; 89JP-000121405.
XX (GREC ) GREEN CROSS CORP.
XX Matsuda H, Ueda Y, Tamanouchi K;
XX WPI; 1990-233117/31.
XX Urokinase precursor-lipid composite - used as thrombolytic agent, having
XX prolonged half-life in the blood, enhanced bioavailability and improved
XX activity.
XX Claim 3; Fig 1; 11pp; English.
XX By forming a precursor-lipid composite, the half-life of this
XX thrombolytic agent in the blood may be increased, exhibiting improved
XX activity without abnormal acceleration of fibrinolytic activity. Compound
XX is useful as a thrombolytic agent in treatment of cerebral thrombosis,
XX myocardial infarction etc
XX Sequence 411 AA;
SQ Query Match 100.0%; Score 554; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTTHAHRSDALQLGLGKHNCRNPDN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTTHAHRSDALQLGLGKHNCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGKKPSPPEE 96
DB 108 RRRPWCYVQVGLKPLVQECWVHDCADGKKPSPPEE 143
RESULT 39
AAR07904
ID AAR07904 standard; protein; 411 AA.
XX AC AAR07904;
XX DT 21-FEB-1991 (first entry)
XX DE Human pro-urokinase variant.
XX KW Thrombin; fibrin; bleeding; pHR27.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Domain 10..42
XX FT Region /label= Epidermal growth factor (EGF) domain
XX FT Region 10..19
XX FT Region /label= First loop
XX FT Region 20..31
XX FT Active-site /label= Second loop
XX FT Region 27..29
XX FT Region /label= Modified site
XX FT Region 33..42
XX FT Region /label= Third loop
XX PN EP380362-A.
XX PD 22-NOV-1990.
XX PF 18-MAY-1990; 90EP-00109473.
XX PR 18-MAY-1989; 89JP-00126434.
XX PA (GREC ) GREEN CROSS CORP.
XX Tarabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;
XX Arimura H;
XX WPI; 1990-350147/47.
XX DR N-PSDB; AAQ06135.
XX Human pro-urokinase variant - produced by recombinant methods, showing
XX increased half life in blood and high affinity for fibrin.
XX Disclosure; Fig 1; 27pp; English.
XX Modified pro-urokinase has a longer half-life in blood, and dissolves
XX thrombin without causing the spontaneous bleeding associated with
XX urokinase. The modification puts an epidermal growth factor domain into
XX the protein, containing the sequence Asn-X-Ser or Asn-X-Thr where X is
XX any residue. Plasmid pHR27 is disclosed as containing the modified
XX sequence
XX Sequence 411 AA;
SQ Query Match 100.0%; Score 554; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTTHAHRSDALQLGLGKHNCRNPDN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTTHAHRSDALQLGLGKHNCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGKKPSPPEE 96
DB 108 RRRPWCYVQVGLKPLVQECWVHDCADGKKPSPPEE 143
RESULT 40
AAR07902
ID AAR07902 standard; protein; 411 AA.
XX AC AAR07902;
XX DT 21-FEB-1991 (first entry)
XX DE Human pro-urokinase variant.
XX KW Thrombin; fibrin; bleeding; pHR22.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Domain 10..42
XX FT Region /label= Epidermal growth factor (EGF) domain
XX FT Region 10..19
XX FT Region /label= First loop
XX FT Region 20..31
XX FT Active-site /label= Second loop
XX FT Region 22..24
XX FT Region /label= Modified site
XX FT Region 33..42
XX FT Region /label= Third loop
XX PN EP380362-A.
XX PD 22-NOV-1990.
XX PF 18-MAY-1990; 90EP-00109473.
XX PA (GREC ) GREEN CROSS CORP.
XX Tarabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;
XX Arimura H;
XX WPI; 1990-350147/47.
XX DR N-PSDB; AAQ06135.
XX Human pro-urokinase variant - produced by recombinant methods, showing
XX increased half life in blood and high affinity for fibrin.
XX Disclosure; Fig 1; 27pp; English.
XX Modified pro-urokinase has a longer half-life in blood, and dissolves
XX thrombin without causing the spontaneous bleeding associated with
XX urokinase. The modification puts an epidermal growth factor domain into
XX the protein, containing the sequence Asn-X-Ser or Asn-X-Thr where X is
XX any residue. Plasmid pHR27 is disclosed as containing the modified
XX sequence
XX Sequence 411 AA;
SQ Query Match 100.0%; Score 554; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTTHAHRSDALQLGLGKHNCRNPDN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTTHAHRSDALQLGLGKHNCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGKKPSPPEE 96
DB 108 RRRPWCYVQVGLKPLVQECWVHDCADGKKPSPPEE 143
```



```
PR 18-MAY-1989; 89JP-00126434.
XX (GREC ) GREEN CROSS CORP.
PA
XX
XX Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;
PI Arimura H;
XX
XX WPI; 1990-350147/47.
XX N-PSDB; AAQ06133.
XX
XX Human pro-urokinase variant - produced by recombinant methods, showing
PT increased half life in blood and high affinity for fibrin.
XX
XX Disclosure; Fig 1; 27pp; English.
XX
XX Modified pro-urokinase has a longer half-life in blood, and dissolves
CC thrombin without causing the spontaneous bleeding associated with
CC urokinase. The modification puts an epidermal growth factor domain into
CC the protein, containing the sequence Asn-X-Ser or Asn-X-Thr where X is
CC any residue. Plasmid pHR22 is disclosed as containing the modified
CC sequence
XX
XX Sequence 411 AA;
SQ
Query Match 100.0%; Score 554; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRSDALQLGLGKKNYCRNPND 60
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRSDALQLGLGKKNYCRNPND 107
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKPSPPPEE 96
DB 108 RRRPWCYVQGLKPLVQECMVHDCADGKPSPPPEE 143
RESULT 41
AAR07903
ID AAR07903 standard; protein; 411 AA.
XX
XX AAR07903;
XX
XX 21-FEB-1991 (first entry)
XX Human pro-urokinase variant.
XX Thrombin; fibrin; bleeding; pHR24.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 10..42 /label= Epidermal growth factor (EGF) domain
XX Region 10..19 /label= First loop
XX Region 20..31 /label= Second loop
XX Active-site 24..26 /label= Modified site
XX Region 33..42 /label= Third loop
XX
XX EP398362-A.
XX
XX 22-NOV-1990.
XX
XX 18-MAY-1990; 90EP-00109473.
XX
XX 18-MAY-1989; 89JP-00126434.
XX (GREC ) GREEN CROSS CORP.
XX
```

```
PI Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;
PI Arimura H;
XX
XX WPI; 1990-350147/47.
XX N-PSDB; AAQ06133.
XX
XX Human pro-urokinase variant - produced by recombinant methods, showing
PT increased half life in blood and high affinity for fibrin.
XX
XX Disclosure; Fig 1; 27pp; English.
XX
XX Modified pro-urokinase has a longer half-life in blood, and dissolves
CC thrombin without causing the spontaneous bleeding associated with
CC urokinase. The modification puts an epidermal growth factor domain into
CC the protein, containing the sequence Asn-X-Ser or Asn-X-Thr where X is
CC any residue. Plasmid pHR24 is disclosed as containing the modified
CC sequence
XX
XX Sequence 411 AA;
SQ
Query Match 100.0%; Score 554; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRSDALQLGLGKKNYCRNPND 60
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRSDALQLGLGKKNYCRNPND 107
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKPSPPPEE 96
DB 108 RRRPWCYVQGLKPLVQECMVHDCADGKPSPPPEE 143
RESULT 42
AAR05117
ID AAR05117 standard; protein; 411 AA.
XX
XX AAR05117;
XX
XX 25-MAR-2003 (revised)
XX 04-OCT-1990 (first entry)
XX
XX UK-S3 as encoded by PUKS3.
XX
XX Urokinase; glycosylation.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 153 /label= synthetic mutation
XX /note= "old seq (Leu)"
XX
XX Misc-difference 155 /label= synthetic mutation
XX /note= "old seq (Pro)"
XX
XX EP370205-A.
XX
XX 30-MAY-1990.
XX
XX 28-SEP-1989; 89EP-00117981.
XX
XX 29-SEP-1988; 88JP-00245705.
XX
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX
XX Sasaki K, Nishi T, Yasumura S, Sato M, Itoh S;
XX
XX WPI; 1990-165029/22.
XX N-PSDB; AAQ04486.
XX
XX Polypeptide(s) with added carbohydrate chains - formed by modification of
PT aminoacid sequence, used to improve physio:chemical properties and/or
```

PT activities.

PS Disclosure; Page ?; 30pp; English.

XX The polypeptide is a deriv. of mature urokinase, designated UK-S3 which

CC has 2 amino acid substns. which result in an N-linked glycosylation site

CC giving the new protein improved stability and activity. See also AAR05113

CC -17. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-

CC 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 411 AA;

XX Query Match 100.0%; Score 554; DB 2; Length 411;

XX Best Local Similarity 100.0%; Pred. No. 2e-44;

XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGREPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNP DN 60

DB 48 KTCYEGNGHFYRGKASTDTMGREPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNP DN 107

QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGKKPSPPEE 96

DB 108 RRRPWCYVQVGLKPLVQECWVHDCADGKKPSPPEE 143

RESULT 43

AAW13634

ID AAW13634 standard; protein; 411 AA.

XX AC AAW13634;

XX 04-JUN-1997 (first entry)

XX Human native prourokinase.

XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;

XX epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.

XX Homo sapiens

XX Key Location/Qualifiers

XX Domain 10..42

XX /label= EGF domain

XX /note= "in claimed variants, at least part of the EGF

XX domain is deleted (see comments)"

XX Region 10..19

XX /label= first_loop

XX Region 20..31

XX /label= second_loop

XX Region 33..42

XX /label= third_loop

XX EP398361-A.

XX 22-NOV-1990.

XX 18-MAY-1990; 90EP-00109472.

XX 18-MAY-1989; 89JP-00126433.

XX 22-FEB-1990; 90JP-00042020.

XX (GREC) GREEN CROSS CORP.

XX Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;

XX Airmura H;

XX WPI; 1990-350146/47.

XX N-PSDB; AAT61671.

XX Human pro-urokinase variants - deficient in loop regions of epidermal

XX growth factor, showing long blood half-life, as fibrinolytic agent.

XX Disclosure; Fig 1; 22pp; English.

XX Query Match 100.0%; Score 554; DB 2; Length 411;

XX Best Local Similarity 100.0%; Pred. No. 2e-44;

XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Sequence 411 AA;

XX Query Match 100.0%; Score 554; DB 2; Length 411;

XX Best Local Similarity 100.0%; Pred. No. 2e-44;

XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGREPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNP DN 60

DB 48 KTCYEGNGHFYRGKASTDTMGREPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNP DN 107

QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGKKPSPPEE 96

DB 108 RRRPWCYVQVGLKPLVQECWVHDCADGKKPSPPEE 143

RESULT 44

AAW10057

ID AAR10057 standard; protein; 411 AA.

XX AC AAR10057;

XX 18-MAR-1991 (first entry)

XX Pro-urokinase derivative UK-T4 with Ser(153) and Thr(155).

XX pro-urokinase; UK-T4; plasminogen activator; myocardial infarction;

XX cerebral thrombosis.

XX Homo sapiens.

XX EP405285-A.

XX 02-JAN-1991.

XX 18-JUN-1990; 90EP-00111471.

XX 19-JUN-1989; 89JP-00156302.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Yasamura S, Nishi T, Ito S;

XX WPI; 1991-008678/02.

XX N-PSDB; AAQ10169.

XX New plasminogen activator almost identical to natural prourokinase - is

XX thrombin resistant and used for prophylaxis-treatment of cerebral

XX thrombosis or myocardial infarction.

XX Disclosure; Page 8; 84pp; English.

XX UK-T4 is one example of a plasminogen activator which differs from

XX natural human pro-urokinase at positions 153 and 155 (Ileu substituted by

XX Ser; Pro substituted by Thr, respectively). The derivative has decreased

XX susceptibility to thrombin compared to natural type pro-UK and higher

XX specific activity. See also AAQ10168 and AAQ10170

XX Sequence 411 AA;

XX Query Match 100.0%; Score 554; DB 2; Length 411;

XX Best Local Similarity 100.0%; Pred. No. 2e-44;

XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPNSATVLQOTYHAHRSDALQGLGKHNCRPN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPNSATVLQOTYHAHRSDALQGLGKHNCRPN 107

QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
DB 108 RRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 143

RESULT 45
AAR10058
ID AAR10058 standard; protein; 411 AA.
XX
AC AAR10058;
XX
DT 18-MAR-1991 (first entry)
XX
DE Pro-urokinase derivative UK-S3 with Asn(153) and Thr(155).
XX
KW pro-urokinase; UK-S3; plasminogen activator; myocardial infarction;
KW cerebral thrombosis.
XX
OS Homo sapiens.
XX
PN EP405285-A.
XX
PD 02-JAN-1991.
XX
PF 18-JUN-1990; 90EP-00111471.
XX
PR 19-JUN-1989; 89JP-00156302.
XX
PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI Yasamura S, Nishi T, Ito S;
XX
DR WPI; 1991-008678/02.
DR N-PSDB; AAQ10170.
XX
PT New plasminogen activator almost identical to natural pro-urokinase - is
PT thrombin resistant and used for prophylaxis-treatment of cerebral
PT thrombosis or myocardial infarction.
XX
PS Disclosure; Page 9; 84pp; English.
XX
CC UK-S3 is one example of a plasminogen activator which differs from
CC natural human pro-urokinase at positions 153 and 155. (Leu substituted by
CC Asn; Pro substituted by Thr, respectively). The derivative has decreased
CC susceptibility to thrombin compared to natural type pro-UK and higher
CC specific activity. See also AAQ10168 and AAQ10169
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 554; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPNSATVLQOTYHAHRSDALQGLGKHNCRPN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPNSATVLQOTYHAHRSDALQGLGKHNCRPN 107

QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
DB 108 RRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 143

RESULT 46
AAR47956
ID AAR47956 standard; protein; 411 AA.
XX
AC AAR47956;
XX

```

```

DT 10-FEB-1994 (first entry)
XX
DE PUK G16A G17A.
XX
KW Pre-urokinase; thrombolytic; blood; plasmid; mutant; PUK.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Domain 10..49
FT FT /label= EGF
FT Region 10..19
FT FT /label= Loop_1
FT Region 20..31
FT FT /label= Loop_2
FT Region 33..42
FT FT /label= Loop_3
XX
PN JP05192142-A.
XX
PD 03-AUG-1993.
XX
PF 20-JAN-1992; 92JP-00030178.
XX
PR 20-JAN-1992; 92JP-00030178.
XX
PA (GREC ) GREEN CROSS CORP.
XX
DR WPI; 1993-277461/35.
XX
PT Mutant human pre-urokinase - by replacing specified aminoacid(s) in N-
PT terminal for providing longer half-life in blood and higher thrombolytic
PT ability.
XX
PS Disclosure; Page 3; 26pp; Japanese.
XX
CC Human pre-urokinase can be modified by replacing: Ser, Asn, Pro, Gly, or
CC Tyr in the region ranging from residue 10 to 60 in the N-terminal of the
CC human PUK by Thr, Pro or Ala. These mutants have a longer half-life in
CC the blood and a higher thrombolytic ability. For examples see (AAR47956-
CC R47960)
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 554; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPNSATVLQOTYHAHRSDALQGLGKHNCRPN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPNSATVLQOTYHAHRSDALQGLGKHNCRPN 107

QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
DB 108 RRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 143

RESULT 47
AAR47966
ID AAR47966 standard; protein; 411 AA.
XX
AC AAR47966;
XX
DT 10-FEB-1994 (first entry)
XX
DE PUK G16A.
XX
KW Pre-urokinase; thrombolytic; blood; plasmid; mutant; PUK.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Domain 10..49

```

```
FT      /label= EGF
FT      10..19
FT      /label= Loop_1
FT      20..31
FT      /label= Loop_2
FT      33..42
FT      /label= Loop_3
XX      JP05192142-A
XX      PD
XX      03-AUG-1993
XX      PF
XX      20-JAN-1992; 92JP-00030178.
XX      PR
XX      20-JAN-1992; 92JP-00030178.
XX      PA
XX      (GREC ) GREEN CROSS CORP.
XX      DR
XX      WPI; 1993-277461/35.
XX      CC
XX      Human pre-urokinase - by replacing specified aminoacid(s) in N-
XX      PT terminal for providing longer half-life in blood and higher thrombolytic
XX      PT ability.
XX      PS
XX      Disclosure; Page 3; 26pp; Japanese.
XX      CC
XX      Human pre-urokinase can be modified by replacing: Ser, Asn, Pro, Gly, or
XX      CC Tyr in the region ranging from residue 10 to 60 in the N-terminal of the
XX      CC human PUK by Thr, Pro or Ala. These mutants have a longer half-life in
XX      CC the blood and a higher thrombolytic ability. For examples see (AAR47956-
XX      CC R47960)
XX      XX
XX      Sequence 411 AA;
XX      QY
XX      Query Match 100.0%; Score 554; DB 2; Length 411;
XX      Best Local Similarity 100.0%; Pred. No. 2e-44;
XX      Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX      QY 1 KTCYEGNGHFGKASTDTMGRCPLPWSATVLCQTYHAHRSDALQLGLGKHNYCRNPDN 60
XX      Db 48 KTCYEGNGHFGKASTDTMGRCPLPWSATVLCQTYHAHRSDALQLGLGKHNYCRNPDN 107
XX      QY 61 RRRPWCYVQGLKPLVQECNVHDCADGKKPSPPEE 96
XX      Db 108 RRRPWCYVQGLKPLVQECNVHDCADGKKPSPPEE 143
XX      RESULT 48
XX      ID AAR47962 standard; protein; 411 AA.
XX      AC AAR47962;
XX      XX
XX      10-FEB-1994 (first entry)
XX      DE
XX      PUK G38A G39A.
XX      KW
XX      Pre-urokinase; thrombolytic; blood; plasmid; mutant; PUK.
XX      OS
XX      Homo sapiens.
XX      FH
XX      Key
XX      Region
XX      Location/Qualifiers
XX      10..49
XX      /label= EGF
XX      10..19
XX      /label= Loop_1
XX      20..31
XX      /label= Loop_2
XX      33..42
XX      /label= Loop_3
XX      FT
XX      Misc-difference 64
XX      FT
XX      Misc-difference /note= "Thr encoded by AGT (sic)"
XX      FT
XX      Misc-difference /note= "Thr encoded by TAC (sic)"
XX      PN
XX      JP05192142-A.
XX      PD
XX      03-AUG-1993.
XX      XX
```

```
FT      /label= EGF
FT      10..19
FT      /label= Loop_1
FT      20..31
FT      /label= Loop_2
FT      33..42
FT      /label= Loop_3
XX      JP05192142-A
XX      PD
XX      03-AUG-1993
XX      PF
XX      20-JAN-1992; 92JP-00030178.
XX      PR
XX      20-JAN-1992; 92JP-00030178.
XX      PA
XX      (GREC ) GREEN CROSS CORP.
XX      DR
XX      WPI; 1993-277461/35.
XX      CC
XX      Human pre-urokinase - by replacing specified aminoacid(s) in N-
XX      PT terminal for providing longer half-life in blood and higher thrombolytic
XX      PT ability.
XX      PS
XX      Disclosure; Page 3; 26pp; Japanese.
XX      CC
XX      Human pre-urokinase can be modified by replacing: Ser, Asn, Pro, Gly, or
XX      CC Tyr in the region ranging from residue 10 to 60 in the N-terminal of the
XX      CC human PUK by Thr, Pro or Ala. These mutants have a longer half-life in
XX      CC the blood and a higher thrombolytic ability. For examples see (AAR47956-
XX      CC R47960)
XX      XX
XX      Sequence 411 AA;
XX      QY
XX      Query Match 100.0%; Score 554; DB 2; Length 411;
XX      Best Local Similarity 100.0%; Pred. No. 2e-44;
XX      Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX      QY 1 KTCYEGNGHFGKASTDTMGRCPLPWSATVLCQTYHAHRSDALQLGLGKHNYCRNPDN 60
XX      Db 48 KTCYEGNGHFGKASTDTMGRCPLPWSATVLCQTYHAHRSDALQLGLGKHNYCRNPDN 107
XX      QY 61 RRRPWCYVQGLKPLVQECNVHDCADGKKPSPPEE 96
XX      Db 108 RRRPWCYVQGLKPLVQECNVHDCADGKKPSPPEE 143
XX      RESULT 48
XX      ID AAR47962 standard; protein; 411 AA.
XX      AC AAR47962;
XX      XX
XX      10-FEB-1994 (first entry)
XX      DE
XX      PUK G38A G39A.
XX      KW
XX      Pre-urokinase; thrombolytic; blood; plasmid; mutant; PUK.
XX      OS
XX      Homo sapiens.
XX      FH
XX      Key
XX      Region
XX      Location/Qualifiers
XX      10..49
XX      /label= EGF
XX      10..19
XX      /label= Loop_1
XX      20..31
XX      /label= Loop_2
XX      33..42
XX      /label= Loop_3
XX      FT
XX      Misc-difference 64
XX      FT
XX      Misc-difference /note= "Thr encoded by AGT (sic)"
XX      FT
XX      Misc-difference /note= "Thr encoded by TAC (sic)"
XX      PN
XX      JP05192142-A.
XX      PD
XX      03-AUG-1993.
XX      XX
```

PF 20-JAN-1992; 92JP-00030178.
XX PR 20-JAN-1992; 92JP-00030178.
XX PA (GREC) GREEN CROSS CORP.
XX DR WPI; 1993-277461/35.
XX DR N-PSDB; AAQ48228.
XX PT Mutant human pre:urokinase - by replacing specified aminoacid(s) in N-
PT terminal for providing longer half-life in blood and higher thrombolytic
XX ability.
XX PS Claim 1; Page 14-16; 26pp; Japanese.
XX CC Human pre-urokinase can be modified by replacing: Ser, Asn, Pro, Gly, or
CC Tyr in the region ranging from residue 10 to 60 in the N-terminal of the
CC human PUK by Thr, Pro or Ala. These mutants have a longer half-life in
CC the blood and a higher thrombolytic ability. For examples see (AAR47956-
CC R47960)
XX SQ Sequence 411 AA;
Query Match 100.0%; Score 554; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. NO. 2e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNCRNPDN 60
Db 48 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 143
RESULT 50
AAR47965
ID AAR47965 standard; protein; 411 AA.
XX AC AAR47965;
XX DT 10-FEB-1994 (first entry)
XX DE PUK N32P G38K G39K.
XX KW Pre-urokinase; thrombolytic; blood; plasmid; mutant; PUK.
XX OS Homo sapiens
XX FH Key
FT Domain 10..49 Location/Qualifiers
FT /label= EGF
FT Region 10..19
FT /label= Loop_1
FT Region 20..31
FT /label= Loop_2
FT Region 33..42
FT /label= Loop_3
XX JP05192142-A.
XX PN 03-AUG-1993.
XX PD 20-JAN-1992; 92JP-00030178.
XX PF 20-JAN-1992; 92JP-00030178.
XX PR (GREC) GREEN CROSS CORP.
XX PA WPI; 1993-277461/35.
XX DR Mutant human pre:urokinase - by replacing specified aminoacid(s) in N-
XX PT terminal for providing longer half-life in blood and higher thrombolytic
XX ability.

PT terminal for providing longer half-life in blood and higher thrombolytic
XX ability.
XX PS Disclosure; Page 3; 26pp; Japanese.
XX CC Human pre-urokinase can be modified by replacing: Ser, Asn, Pro, Gly, or
CC Tyr in the region ranging from residue 10 to 60 in the N-terminal of the
CC human PUK by Thr, Pro or Ala. These mutants have a longer half-life in
CC the blood and a higher thrombolytic ability. For examples see (AAR47956-
CC R47960)
XX SQ Sequence 411 AA;
Query Match 100.0%; Score 554; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. NO. 2e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNCRNPDN 60
Db 48 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 143
Search completed: May 25, 2004, 14:53:02
Job time : 24.4741 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:53:05 ; Search time 17.5004 Seconds
(without alignments)
1530.046 Million cell updates/sec

Title: US-09-880-503-9

Perfect score: 554

Sequence: 1 KTCYEGNGHFYRGKASTDTM.....QECMVHDCADGKKRSPPEE 96

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1149313 seqs, 278921704 residues

Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	554	100.0	96	9	US-09-880-503-9
2	554	100.0	143	9	US-09-880-503-8
3	554	100.0	337	14	US-10-106-698-6366
4	554	100.0	337	15	US-10-264-049-2927
5	554	100.0	411	9	US-09-880-503-3
6	554	100.0	431	9	US-09-264-4688-1
7	554	100.0	431	12	US-10-411-037-34
8	554	100.0	431	12	US-10-411-028-34
9	554	100.0	431	13	US-10-076-421-2
10	554	100.0	431	14	US-10-171-311-184
11	554	100.0	431	14	US-10-193-656-4
12	554	100.0	431	14	US-10-301-822-161
13	554	100.0	431	14	US-10-131-985-21
14	554	100.0	431	15	US-10-295-027-414
15	554	100.0	431	15	US-10-295-027-1275

16	554	100.0	431	16	US-10-410-962-34	Sequence 34, Appl
17	554	100.0	431	16	US-10-411-049-34	Sequence 34, Appl
18	554	100.0	437	12	US-10-087-192-594	Sequence 594, App
19	552	99.6	431	14	US-10-247-671-149	Sequence 149, App
20	545	98.4	431	12	US-10-282-174-562	Sequence 562, App
21	544	98.2	411	15	US-10-407-821-2	Sequence 2, Appl
22	543	98.0	445	15	US-10-360-101-266	Sequence 266, App
23	510	92.1	88	9	US-09-880-503-1	Sequence 1, Appl
24	510	92.1	135	9	US-09-880-503-4	Sequence 4, Appl
25	510	92.1	138	9	US-09-884-186-12	Sequence 12, Appl
26	510	92.1	138	14	US-10-237-667-12	Sequence 12, Appl
27	510	92.1	138	14	US-10-237-708-12	Sequence 12, Appl
28	510	92.1	138	14	US-10-237-866-12	Sequence 12, Appl
29	510	92.1	138	14	US-10-237-871-12	Sequence 12, Appl
30	510	92.1	138	14	US-10-237-624-12	Sequence 12, Appl
31	510	92.1	138	16	US-10-702-536-12	Sequence 12, Appl
32	510	92.1	138	16	US-10-702-636-12	Sequence 12, Appl
33	510	92.1	403	9	US-09-880-503-6	Sequence 6, Appl
34	504	91.0	322	12	US-10-424-999-21	Sequence 21, Appl
35	504	91.0	322	12	US-10-425-000-41	Sequence 41, Appl
36	504	91.0	322	15	US-10-233-675A-21	Sequence 21, Appl
37	500	90.3	687	12	US-10-424-999-17	Sequence 17, Appl
38	500	90.3	687	12	US-10-425-000-37	Sequence 37, Appl
39	500	90.3	687	13	US-10-233-675A-17	Sequence 17, Appl
40	499	90.1	86	12	US-10-424-999-5	Sequence 5, Appl
41	499	90.1	86	12	US-10-424-999-62	Sequence 62, Appl
42	499	90.1	86	12	US-10-425-000-97	Sequence 97, Appl
43	499	90.1	86	15	US-10-233-675A-5	Sequence 5, Appl
44	499	90.1	86	15	US-10-233-675A-22	Sequence 22, Appl
45	499	90.1	87	12	US-10-424-999-9	Sequence 9, Appl
46	499	90.1	87	12	US-10-425-000-29	Sequence 29, Appl
47	499	90.1	87	15	US-10-233-675A-9	Sequence 9, Appl
48	499	90.1	91	12	US-10-424-999-37	Sequence 37, Appl
49	499	90.1	221	12	US-10-424-999-36	Sequence 36, Appl
50	494	89.2	86	12	US-10-424-999-1	Sequence 1, Appl
51	494	89.2	86	15	US-10-233-675A-1	Sequence 1, Appl
52	494	89.2	87	12	US-10-424-999-10	Sequence 10, Appl
53	494	89.2	87	12	US-10-425-000-30	Sequence 30, Appl
54	494	89.2	87	15	US-10-233-675A-10	Sequence 10, Appl
55	494	89.2	91	12	US-10-424-999-35	Sequence 35, Appl
56	494	89.2	221	12	US-10-424-999-34	Sequence 34, Appl
57	494	89.2	322	12	US-10-424-999-20	Sequence 20, Appl
58	494	89.2	322	12	US-10-425-000-40	Sequence 40, Appl
59	494	89.2	322	15	US-10-233-675A-20	Sequence 20, Appl
60	494	89.2	672	12	US-10-424-999-15	Sequence 15, Appl
61	494	89.2	672	12	US-10-425-000-35	Sequence 35, Appl
62	494	89.2	672	15	US-10-233-675A-15	Sequence 15, Appl
63	494	89.2	674	12	US-10-424-999-14	Sequence 14, Appl
64	494	89.2	674	12	US-10-425-000-34	Sequence 34, Appl
65	494	89.2	674	15	US-10-233-675A-14	Sequence 14, Appl
66	494	89.2	688	12	US-10-424-999-18	Sequence 18, Appl
67	494	89.2	688	12	US-10-425-000-38	Sequence 38, Appl
68	494	89.2	688	15	US-10-233-675A-18	Sequence 18, Appl
69	494	89.2	689	12	US-10-424-999-13	Sequence 13, Appl
70	494	89.2	689	12	US-10-425-000-33	Sequence 33, Appl
71	494	89.2	689	15	US-10-233-675A-13	Sequence 13, Appl
72	493	89.0	650	15	US-10-401-077-1	Sequence 1, Appl
73	489	88.3	86	12	US-10-424-999-7	Sequence 7, Appl
74	489	88.3	86	13	US-10-233-675A-7	Sequence 7, Appl
75	483	87.2	86	15	US-10-233-675A-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-09-880-503-9
; Sequence 9, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND

```
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-880-503-9

Query Match      100.0%; Score 554; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.4e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQGLGKHN YCRNPDN 60
DB 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQGLGKHN YCRNPDN 60
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 96
DB 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 96

RESULT 2
US-09-880-503-8
; Sequence 8, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-880-503-8

Query Match      100.0%; Score 554; DB 9; Length 143;
Best Local Similarity 100.0%; Pred. No. 2.1e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQGLGKHN YCRNPDN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQGLGKHN YCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 96
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 143

RESULT 3
US-10-106-698-6266
; Sequence 6266, Application US/10106698
; Patent No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA00591
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
```

```
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patentin Ver. 3.0
; SEQ ID NO 6266
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-106-698-6266

Query Match      100.0%; Score 554; DB 14; Length 337;
Best Local Similarity 100.0%; Pred. No. 5.5e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQGLGKHN YCRNPDN 60
DB 74 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQGLGKHN YCRNPDN 133
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 96
DB 134 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 169

RESULT 4
US-10-264-049-2927
; Sequence 2927, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al
; APPLICANT: CINES, Douglas B
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: Patentin Ver. 3.1
; SEQ ID NO 2927
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-264-049-2927

Query Match      100.0%; Score 554; DB 15; Length 337;
Best Local Similarity 100.0%; Pred. No. 5.5e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQGLGKHN YCRNPDN 60
DB 74 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQGLGKHN YCRNPDN 133
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 96
DB 134 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 169

RESULT 5
US-09-880-503-3
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
```

```

; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-3

Query Match          100.0%; Score 554; DB 9; Length 411;
Best Local Similarity 100.0%; Pred. No. 6.8e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTYYAHRS DALQLGLGKHNYCRNPDN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTYYAHRS DALQLGLGKHNYCRNPDN 107

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 143

RESULT 6
US-09-264-468B-1
; Sequence 1, Application US/09264468B
; Patent No. US20020106778A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.P1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
; OTHER INFORMATION: Leader sequence
; NAME/KEY: VARIANT
; LOCATION: (279)...(279)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (302)...(302)
; OTHER INFORMATION: Xaa = any amino acid
US-09-264-468B-1

Query Match          100.0%; Score 554; DB 9; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.2e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTYYAHRS DALQLGLGKHNYCRNPDN 60
DB 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTYYAHRS DALQLGLGKHNYCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 163
```

```

RESULT 7
US-10-411-037-34
; Sequence 34, Application US/10411037
; Publication No. US20040043446A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bove, Caryn
; TITLE OF INVENTION: ALPHA GALACTOSIDASE A: REMODELING AND GLYCOCONJUGATION OF ALPHA
; FILE REFERENCE: 040853-01-5082
; CURRENT APPLICATION NUMBER: US/10/411,037
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-037-34

Query Match          100.0%; Score 554; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.2e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTYYAHRS DALQLGLGKHNYCRNPDN 60
DB 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTYYAHRS DALQLGLGKHNYCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 163

RESULT 8
US-10-411-026-34
; Sequence 34, Application US/10411026
; Publication No. US20040063911A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: PROTEIN REMODELING METHODS AND PROTEINS/PEPTIDES PRODUCED BY THE
; FILE REFERENCE: 040853-01-5053
; CURRENT APPLICATION NUMBER: US/10/411,026
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
```



```
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-026-34

Query Match      100.0%; Score 554; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.2e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 127

Qy 61 RRRPWCYVQVGLKPLVQECWVHDCADGKSPSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECWVHDCADGKSPSPPEE 163

RESULT 9
US-10-076-421-2
; Sequence 2, Application US/10076421
; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, NAOKO
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAYAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-2

Query Match      100.0%; Score 554; DB 13; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.2e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 127

Qy 61 RRRPWCYVQVGLKPLVQECWVHDCADGKSPSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECWVHDCADGKSPSPPEE 163

RESULT 10
US-10-171-311-184
; Sequence 184, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
```

```
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Ganavarapu, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-184

Query Match      100.0%; Score 554; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.2e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKKNYCRNPDN 127

Qy 61 RRRPWCYVQVGLKPLVQECWVHDCADGKSPSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECWVHDCADGKSPSPPEE 163

RESULT 11
US-10-193-656-4
; Sequence 4, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00749
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)..(431)
US-10-193-656-4

Query Match      100.0%; Score 554; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.2e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOITYHAHRS DALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOITYHAHRS DALQLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 128 RRRPWCYVQGLKPLVQECMVHDCADGKKPSSPPEE 163

RESULT 12
US-10-301-822-161
; Sequence 151, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MEM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match 100.0%; Score 554; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.2e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOITYHAHRS DALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOITYHAHRS DALQLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 128 RRRPWCYVQGLKPLVQECMVHDCADGKKPSSPPEE 163

RESULT 13
US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication No. US2003019940A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Occleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-985-21

Query Match 100.0%; Score 554; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.2e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOITYHAHRS DALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOITYHAHRS DALQLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 128 RRRPWCYVQGLKPLVQECMVHDCADGKKPSSPPEE 163

RESULT 14
US-10-295-027-414
; Sequence 414, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afari, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 414
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-414

Query Match 100.0%; Score 554; DB 15; Length 431;

Best Local Similarity 100.0%; Pred. No. 7.2e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 127
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 163

RESULT 15
US-10-295-027-1275
; Sequence 1275, Application US/10295027
; Publication No. US2003023350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1275
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1275

Query Match 100.0%; Score 554; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.2e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 127
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 163

RESULT 16
US-10-410-962-34
; Sequence 34, Application US/10410962
; Publication No. US20040077836A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bove, Caryn
; TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR: REMODELING AND
; TITLE OF INVENTION: GLYCOCONJUGATION OF G-CSF
; FILE REFERENCE: 040853-01-5054
; CURRENT APPLICATION NUMBER: US/10/410,962
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 34
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-410-962-34
Query Match 100.0%; Score 554; DB 16; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.2e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 127
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 163

RESULT 17
US-10-411-049-34
; Sequence 34, Application US/10411049
; Publication No. US20040082026A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bove, Caryn
; TITLE OF INVENTION: INTERFERON ALPHA: REMODELING AND GLYCOCONJUGATION OF INTERFERON
; TITLE OF INVENTION: ALPHA
; FILE REFERENCE: 040853-01-5055
; CURRENT APPLICATION NUMBER: US/10/411,049
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692

```
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 34
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-049-34

Query Match      100.0%; Score 554; DB 16; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.2e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQGLGKHNCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQGLGKHNCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 163

RESULT 18
US-10-087-192-594
; Sequence 594, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 594
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-594

Query Match      100.0%; Score 554; DB 12; Length 437;
Best Local Similarity 100.0%; Pred. No. 7.3e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQGLGKHNCRNPDN 60
Db 74 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQGLGKHNCRNPDN 133

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 96
Db 134 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 169

RESULT 19
US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match      99.6%; Score 552; DB 14; Length 431;
Best Local Similarity 99.0%; Pred. No. 1.2e-55;
Matches 95; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQGLGKHNCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQGLGKHNCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 163

RESULT 20
US-10-282-174-562
; Sequence 562, Application US/10282174
; Publication No. US20030224380A1
; GENERAL INFORMATION:
; APPLICANT: Becker, Kenneth David
; APPLICANT: Velicelebi, Gonul
; APPLICANT: Elliot, Kathryn J.
; APPLICANT: Wang, Xin
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Bertram, Lars
; APPLICANT: Saunders, Alistair J.
; APPLICANT: Mullin, Kristina M.
; APPLICANT: Sampson, Andrew Johnson
; APPLICANT: Blacker, Deborah Lynne
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
; TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
; TITLE OF INVENTION: NEURODEGENERATIVE DISEASES
; FILE REFERENCE: 37481-3308
; CURRENT APPLICATION NUMBER: US/10/282,174
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US 60/339,525
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 60/338,010
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/336,929
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/338,363
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/337,052
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/368,919
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 564
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 562
; LENGTH: 431
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 15, 58, 141, 214, 231, 274, 366
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-282-174-562

Query Match          98.4%; Score 545; DB 12; Length 431;
Best Local Similarity 99.0%; Pred. No. 8e-55;
Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNCRNPDN 60
DB 58 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNCRNPDN 127
QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPSPPEE 96
DB 128 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPSPPEE 163

RESULT 21
US-10-407-821-2
; Sequence 2, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
; FILE REFERENCE: UTSN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-2

Query Match          98.2%; Score 544; DB 15; Length 411;
Best Local Similarity 99.0%; Pred. No. 9.9e-55;
Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNCRNPDN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPSPPEE 96
DB 108 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPSPPEE 143

RESULT 22
US-10-360-101-266
; Sequence 266, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 266
```

```
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of urokinase
US-10-360-101-266

Query Match          98.0%; Score 543; DB 15; Length 445;
Best Local Similarity 98.0%; Pred. No. 1.4e-54;
Matches 96; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTY--HAHRSALQLGLGKHNCRNP 58
DB 70 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYFTHAHRSDALQLGLGKHNCRNP 129
QY 59 DNRERPWCYVQVGLKPLVQECMWHDCADGKPKSPSPPEE 96
DB 130 DNRERPWCYVQVGLKPLVQECMWHDCADGKPKSPSPPEE 167

RESULT 23
US-09-880-503-1
; Sequence 1, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-1

Query Match          92.1%; Score 510; DB 9; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.6e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNCRNPDN 60
DB 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNCRNPDN 60
QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88
DB 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88

RESULT 24
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-4

Query Match          92.1%; Score 510; DB 9; Length 135;
Best Local Similarity 100.0%; Pred. No. 2.6e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 107

QY 61 RRPWCYVQVGLKPLVQECWVHDCADGK 88
Db 108 RRPWCYVQVGLKPLVQECWVHDCADGK 135

RESULT 25
US-09-984-186-12
; Sequence 12, Application US/09984186
; Patent No. US20020151011A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12
Query Match          92.1%; Score 510; DB 9; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.7e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 51 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 110

QY 61 RRPWCYVQVGLKPLVQECWVHDCADGK 88
Db 111 RRPWCYVQVGLKPLVQECWVHDCADGK 138

RESULT 26
US-10-237-667-12
; Sequence 12, Application US/10237667
; Publication No. US20030022306A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,667
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-667-12
Query Match          92.1%; Score 510; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.7e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
```

```

Db      51 KTCYEGNGHFRYGKASTDTMGRCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 110
QY      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 27
US-10-237-708-12
; Sequence 12, Application US/10237708
; Publication No. US20030036170A1
; GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Gultton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
;
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
;
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,708
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph. D. Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12

Query Match 92.1%; Score 510; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.7e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KTCYEGNGHFRYGKASTDTMGRCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
Db      51 KTCYEGNGHFRYGKASTDTMGRCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 110
QY      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

```

```

Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (patentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,624
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38 619
REFERENCE/DOCKET NUMBER: SF92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-624-12
Query Match 92.1%; Score 510; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.7e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTWGRCLPWSNATVLQOYTHAHRSDALQLGLGKHNYCRPND 60
DB 51 KTCYEGNGHFYRGKASTDTWGRCLPWSNATVLQOYTHAHRSDALQLGLGKHNYCRPND 110
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 111 RRRPKCYVQVGLKPLVQECMVHDCADGK 138
RESULT 31
US-10-702-536-12
Sequence 12, Application US/10702536
Publication No. US20040086976A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guittou, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

```


NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/702,536
FILING DATE: 07-Nov-2003
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELEPHONE: (610) 454-3808
TELEFAX: (610) 454-3839
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-702-536-12
Query Match 92.1%; Score 510; DB 16; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.7e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
US-10-702-536-12
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPDN 60
DB 51 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPDN 110
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGK 88
DB 111 RRRPWCYVQGLKPLVQECMVHDCADGK 138
RESULT 32
US-10-702-636-12
Sequence 12, Application US/10702636
Publication No. US20040086977A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville

STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/702,636
FILING DATE: 06-Nov-2003
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-702-636-12
Query Match 92.1%; Score 510; DB 16; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.7e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPDN 60
DB 51 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPDN 110
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGK 88
DB 111 RRRPWCYVQGLKPLVQECMVHDCADGK 138
RESULT 33
US-09-880-503-6
Sequence 6, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
TITLE OF INVENTION: TISSUE CONTRACTILITY
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 403
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-6
Query Match 92.1%; Score 510; DB 9; Length 403;

```
Best Local Similarity 100.0%; Pred. No. 8.7e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 107

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 34
US-10-424-999-21
; Sequence 21, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-21

Query Match
Best Local Similarity 91.0%; Score 504; DB 12; Length 322;
Matches 87; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 61

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSP 93
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADRLPRGP 94

RESULT 35
US-10-425-000-41
; Sequence 41, Application US/10425000
; Publication No. US2004005277A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; TITLE OF INVENTION: Kringie Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 41
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
```

```
US-10-425-000-41

Query Match
Best Local Similarity 91.0%; Score 504; DB 12; Length 322;
Matches 87; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 61

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSP 93
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADRLPRGP 94

RESULT 36
US-10-233-675A-21
; Sequence 21, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods fo
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion protein human abrogen
US-10-233-675A-21

Query Match
Best Local Similarity 91.0%; Score 504; DB 15; Length 322;
Matches 87; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 61

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSP 93
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADRLPRGP 94

RESULT 37
US-10-424-999-17
; Sequence 17, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods fo
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 687
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-17

Query Match          90.3%; Score 500; DB 12; Length 687;
Best Local Similarity 98.9%; Pred. No. 2.3e-49;
Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 60
   |||
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 61
   |||

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
   |||
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADG 88

RESULT 38
US-10-425-000-37
; Sequence 37, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-37

Query Match          90.3%; Score 500; DB 12; Length 687;
Best Local Similarity 98.9%; Pred. No. 2.3e-49;
Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 60
   |||
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 61
   |||

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
   |||
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADG 88

RESULT 39
US-10-233-675A-17
; Sequence 17, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 17
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion protein human abrogen
US-10-233-675A-17

Query Match          90.3%; Score 500; DB 15; Length 687;
Best Local Similarity 98.9%; Pred. No. 2.3e-49;
Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 60
   |||
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 61
   |||

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
   |||
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADG 88

RESULT 40
US-10-424-999-5
; Sequence 5, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human abrogen
US-10-424-999-5

Query Match          90.1%; Score 499; DB 12; Length 86;
Best Local Similarity 100.0%; Pred. No. 3e-50;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 60
   |||
Db 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRNPDN 60
   |||

QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
   |||
Db 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 41
US-10-424-999-62
; Sequence 62, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
```

```
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 62
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human Kringle domain ATF-Kringle (Abrogen)
US-10-424-999-62

Query Match          90.1%; Score 499; DB 12; Length 86;
Best Local Similarity 100.0%; Pred. No. 3e-50;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNCRPN 60
DB 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNCRPN 60
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 42
US-10-425-000-97
; Sequence 97, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 97
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human kringle domain ATF-Kringle (Abrogen)
US-10-425-000-97

Query Match          90.1%; Score 499; DB 12; Length 86;
Best Local Similarity 100.0%; Pred. No. 3e-50;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNCRPN 60
DB 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNCRPN 60
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 43
US-10-233-675A-5
; Sequence 5, Application US/10233675A
; Publication No. US2003028298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Brockstedt, Dirk
; APPLICANT: Fong, Timothy
; TITLE OF INVENTION: Abrogen Polypeptides; Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027
```

```
; CURRENT APPLICATION NUMBER: US/10/233.675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
US-10-233-675A-5

Query Match          90.1%; Score 499; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 3e-50;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNCRPN 60
DB 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNCRPN 60
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 44
US-10-233-675A-22
; Sequence 22, Application US/10233675A
; Publication No. US2003028298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides; Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233.675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fragment of human urokinase plasminogen activator
US-10-233-675A-22

Query Match          90.1%; Score 499; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 3e-50;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNCRPN 60
DB 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNCRPN 60
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 45
US-10-424-999-9
; Sequence 9, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
```

```
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human abrogen as secreted from pME063 (abrogen D43)
US-10-424-999-9

Query Match          90.1%; Score 499; DB 12; Length 87;
Best Local Similarity 100.0%; Pred. No. 3.1e-50;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 61

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
Db 62 RRRPWCYVQVGLKPLVQECMVHDCAD 87

RESULT 46
US-10-425-000-29
; Sequence 29, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human abrogen (D43)
US-10-425-000-29

Query Match          90.1%; Score 499; DB 12; Length 87;
Best Local Similarity 100.0%; Pred. No. 3.1e-50;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 61

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
Db 62 RRRPWCYVQVGLKPLVQECMVHDCAD 87

RESULT 47
US-10-233-675A-9
; Sequence 9, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Nesbit, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived fusion protein
US-10-233-675A-9

Query Match          90.1%; Score 499; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 3.1e-50;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 61

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
Db 62 RRRPWCYVQVGLKPLVQECMVHDCAD 87

RESULT 48
US-10-424-999-37
; Sequence 37, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Abrogen D43
US-10-424-999-37

Query Match          90.1%; Score 499; DB 12; Length 91;
Best Local Similarity 100.0%; Pred. No. 3.2e-50;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 60
Db 6 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 65

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
Db 66 RRRPWCYVQVGLKPLVQECMVHDCAD 91

RESULT 49
US-10-424-999-36
```

Search completed: May 25, 2004, 15:03:49
Job time : 18.5004 secs

```
; Sequence 36, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: TrxA-Abrogen D43 fusion protein
US-10-424-999-36

Query Match          90.1%; Score 499; DB 12; Length 221;
Best Local Similarity 100.0%; Pred. No. 8.6e-50;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
Db 136 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 195

QY 61 RRPWCYVQVGLKPLVQECWVHDCAD 86
Db 196 RRPWCYVQVGLKPLVQECWVHDCAD 221

RESULT 50
US-10-424-999-1
; Sequence 1, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human abrogen N43
US-10-424-999-1

Query Match          89.2%; Score 494; DB 12; Length 86;
Best Local Similarity 98.8%; Pred. No. 1.1e-49;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
Db 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60

QY 61 RRPWCYVQVGLKPLVQECWVHDCAD 86
Db 61 RRPWCYVQVGLKPLVQECWVHDCAD 86
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:48:05 ; Search time 6.6468 Seconds
(without alignments)
745.636 Million cell updates/sec

Title: US-09-880-503-9

Perfect score: 554

Sequence: 1 KTCYEGNGHYRGKASTDTM.....QECMVHDCADGKFPSPPEE 96

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCITUS_COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	554	100.0	200	4	US-09-101-272G-73
2	554	100.0	208	4	US-09-101-272G-98
3	554	100.0	365	1	US-08-093-741-83
4	554	100.0	365	1	US-08-720-012-83
5	554	100.0	393	2	US-08-560-098A-44
6	554	100.0	393	3	US-08-367-024C-24
7	554	100.0	393	3	US-08-967-024C-25
8	554	100.0	411	1	US-08-087-163-1
9	554	100.0	411	1	US-08-286-748B-18
10	554	100.0	411	1	US-08-153-799-18
11	554	100.0	411	4	US-09-403-736-2
12	554	100.0	430	1	US-07-942-157A-3
13	554	100.0	431	4	US-09-101-272G-1
14	554	100.0	431	6	518829-1
15	554	100.0	432	2	US-08-560-098A-47
16	554	98.2	411	3	US-09-181-816-1
17	543	98.0	411	2	US-08-560-098A-48
18	538.5	97.2	430	6	5219569-2
19	530	95.7	157	3	US-08-142-590B-25
20	510	92.1	138	2	US-08-787-689-12
21	510	92.1	138	4	US-09-984-186-12
22	505	91.2	194	4	US-09-101-272G-80
23	505	91.2	201	4	US-09-101-272G-96
24	489	88.3	89	4	US-09-101-272G-62
25	241	43.5	477	2	US-08-560-098A-51
26	226	40.8	527	1	US-07-609-510B-16
27	226	40.8	527	2	US-08-811-949-39

ALIGNMENTS

RESULT 1
US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101.272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

28	226	40.8	527	5	PCT-US91-01025A-2	Sequence 2, Appli
29	226	40.8	527	6	5185259-8	Patent No. 5185259
30	226	40.8	527	6	520913-1	Patent No. 520913
31	226	40.8	546	6	5200340-6	Patent No. 5200340
32	226	40.8	562	2	US-08-811-949-43	Sequence 43, Appl
33	226	40.8	562	2	US-08-560-098A-50	Sequence 50, Appl
34	226	40.8	562	2	US-08-883-795A-38	Sequence 38, Appl
35	226	40.8	562	4	US-09-703-695A-4	Sequence 4, Appli
36	226	40.8	562	6	5185259-3	Patent No. 5185259
37	226	40.8	562	6	5200340-2	Patent No. 5200340
38	226	40.8	562	6	5344773-2	Patent No. 5344773
39	221.5	40.0	356	1	US-08-427-640-8	Sequence 8, Appli
40	221	39.9	437	2	US-08-811-949-49	Sequence 49, Appl
41	221	39.9	437	2	US-08-811-949-51	Sequence 51, Appl
42	221	39.9	437	2	US-08-811-949-55	Sequence 55, Appl
43	221	39.9	437	2	US-08-811-949-57	Sequence 57, Appl
44	220.5	39.8	378	4	US-09-553-498-10	Sequence 10, Appl
45	220.5	39.8	378	4	US-09-618-869-10	Sequence 10, Appl
46	219.5	39.6	389	2	US-08-811-949-67	Sequence 67, Appl
47	218.5	39.4	386	4	US-09-111-977-3	Sequence 3, Appli
48	218.5	39.4	355	1	US-08-137-116-1	Sequence 1, Appli
49	218.5	39.4	355	1	US-08-217-618-1	Sequence 1, Appli
50	218.5	39.4	355	1	US-08-427-640-2	Sequence 2, Appli
51	218.5	39.4	355	1	US-08-217-617A-1	Sequence 1, Appli
52	218.5	39.4	355	1	US-08-217-616-1	Sequence 1, Appli
53	218.5	39.4	355	2	US-08-811-949-45	Sequence 45, Appl
54	218.5	39.4	355	2	US-08-811-949-47	Sequence 47, Appl
55	218.5	39.4	355	2	US-08-811-949-53	Sequence 53, Appl
56	218.5	39.4	355	3	US-08-794-528-1	Sequence 1, Appli
57	218.5	39.4	355	6	5223256-1	Patent No. 5223256
58	218.5	39.4	389	2	US-08-811-949-65	Sequence 65, Appl
59	218.5	39.4	472	2	US-08-811-949-63	Sequence 63, Appl
60	218.5	39.4	562	6	524676-5	Patent No. 524676
61	216.5	39.1	354	2	US-08-811-949-61	Sequence 61, Appl
62	215	38.8	347	2	US-08-811-949-59	Sequence 59, Appl
63	215	38.8	355	2	US-08-811-949-59	Sequence 59, Appl
64	213.5	38.5	355	1	US-08-427-640-6	Sequence 6, Appli
65	213	38.4	83	2	US-08-811-949-2	Sequence 2, Appli
66	204	36.8	356	1	US-08-427-640-4	Sequence 4, Appli
67	194.5	35.1	655	1	US-08-148-910-12	Sequence 12, Appl
68	194.5	35.1	655	1	US-08-448-337A-12	Sequence 12, Appl
69	154	27.8	458	3	US-09-039-609-2	Sequence 2, Appli
70	152	27.4	812	4	US-08-991-761A-7	Sequence 7, Appli
71	149	26.9	810	1	US-08-991-761A-11	Sequence 11, Appl
72	148.5	26.8	711	1	US-08-184-012C-8	Sequence 8, Appli
73	148.5	26.8	711	1	US-08-334-177-2	Sequence 2, Appli
74	148.5	26.8	711	2	US-08-666-082B-1	Sequence 1, Appli
75	148.5	26.8	711	2	US-08-766-382-2	Sequence 2, Appli

OTHER INFORMATION: ATP domain of uPA
US-09-101-272G-73

Query Match 100.0%; Score 554; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 3e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHYRGKASTDTMGRCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 127
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 163

RESULT 2

US-09-101-272G-98
Sequence 98, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: Q50979
CURRENT APPLICATION NUMBER: US/09/101,272G
PRIOR FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patent in version 3.1
SEQ ID NO 98
LENGTH: 208
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: ATFHI-ML chimeric protein
US-09-101-272G-98

Query Match 100.0%; Score 554; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 3.2e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHYRGKASTDTMGRCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
Db 49 KTCYEGNGHYRGKASTDTMGRCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 108
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
Db 109 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 144

RESULT 3

US-08-093-741-83
Sequence 83, Application US/08093741
Patent No. 5681721
GENERAL INFORMATION:
APPLICANT: STEPPENS, GERO J.
APPLICANT: WENDT, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,741
FILING DATE: 20-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-093-741-83

Query Match 100.0%; Score 554; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 6.2e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHYRGKASTDTMGRCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHYRGKASTDTMGRCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 61
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 97

RESULT 4

US-08-720-012-83
Sequence 83, Application US/08720012
Patent No. 5747291
GENERAL INFORMATION:
APPLICANT: STEPPENS, GERO J.
APPLICANT: WENDT, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,012
FILING DATE: 27-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,741
FILING DATE: 20-JUL-1993
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.


```
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/41345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-8800
; TELEFAX: (202)628-8844
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-720-012-83

Query Match 100.0%; Score 554; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 6.2e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60
Db 3 KTCYEGNGHFYRGKASTDTMGPRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 62

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 96
Db 63 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 98

RESULT 6
US-08-967-024C-24
; Sequence 24, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNENDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 42 565.8
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-967-024C-24

Query Match 100.0%; Score 554; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60
Db 3 KTCYEGNGHFYRGKASTDTMGPRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 62

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 96
Db 63 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 98

RESULT 7
US-08-967-024C-25
; Sequence 25, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNENDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 42 565.8
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-967-024C-25

Query Match 100.0%; Score 554; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60
Db 3 KTCYEGNGHFYRGKASTDTMGPRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 62

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 96
Db 63 RRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEE 98

RESULT 8
US-08-560-098A-44
; Sequence 44, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNENDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-44

Query Match 100.0%; Score 554; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

APPLICANT: WNEEDT, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOSCH, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stevenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 100.0%; Score 554; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHYRGKASTDTWGRCLPWNSATVLOQTYHAHRSDALQGLGKHNCRNPDN 60
Db 3 KTCYEGNGHYRGKASTDTWGRCLPWNSATVLOQTYHAHRSDALQGLGKHNCRNPDN 62

Qy 61 RRRPWCYVQGLKPLVQECMVHDCADGKXPSPPEE 96
Db 63 RRRPWCYVQGLKPLVQECMVHDCADGKXPSPPEE 98

RESULT 8
US-08-087-163-1
Sequence 1, Application US/08087163
Patent No. 5472692
GENERAL INFORMATION:
APPLICANT: Liu, Jian-Ning
APPLICANT: Gurewicz, Victor
TITLE OF INVENTION: PRO-UKINASE MUTANTS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)

SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,163
FILING DATE: 07/02/93
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04353/003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
US-08-087-163-1

Query Match 100.0%; Score 554; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 7.2e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHYRGKASTDTWGRCLPWNSATVLOQTYHAHRSDALQGLGKHNCRNPDN 60
Db 48 KTCYEGNGHYRGKASTDTWGRCLPWNSATVLOQTYHAHRSDALQGLGKHNCRNPDN 107

Qy 61 RRRPWCYVQGLKPLVQECMVHDCADGKXPSPPEE 96
Db 108 RRRPWCYVQGLKPLVQECMVHDCADGKXPSPPEE 143

RESULT 9
US-08-286-748B-18
Sequence 18, Application US/08286748B
Patent No. 5759542
GENERAL INFORMATION:
APPLICANT: Victor Gurewicz
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,748B
FILING DATE: August 5, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: J. Peter Fasse
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04547/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906

```
;
;
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-286-748B-18
;
; Query Match 100.0%; Score 554; DB 1; Length 411;
; Best Local Similarity 100.0%; Pred. No. 7.2e-58;
; Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 KTCYEGNGHFGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 60
; DB 48 KTCYEGNGHFGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 107
;
; QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPSPPEE 96
; DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPSPPEE 143
;
; RESULT 10
; US-08-153-799-18
; Sequence 18, Application US/08153799
; Patent No. 5765883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA: PCT/GB90/00650
; APPLICATION NUMBER:
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
US-08-153-799-18
;
; Query Match 100.0%; Score 554; DB 1; Length 411;
; Best Local Similarity 100.0%; Pred. No. 7.2e-58;
; Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 KTCYEGNGHFGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 60
; DB 48 KTCYEGNGHFGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 107
;
; QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPSPPEE 96
; DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPSPPEE 143
;
; RESULT 11
; US-09-403-736-2
; Sequence 2, Application US/09403736
; Patent No. 6638502
; GENERAL INFORMATION:
; APPLICANT: Aventis S.A.
; APPLICANT: Li, Hong
; APPLICANT: LU, He
; APPLICANT: GRISCELLI, Frank
; APPLICANT: OPOLON, Paule
; APPLICANT: SORIA, Claudine
; APPLICANT: RAGOT, Thierry
; APPLICANT: LERAND, Yves
; APPLICANT: SORIA, Jeanette
; APPLICANT: MABILAT, Christelle
; APPLICANT: PERRICAUDET, Michel
; APPLICANT: YEH, Patrice
; TITLE OF INVENTION: Adenovirus-Mediated Intratumoral Delivery Of An Angiogenesis Ant
; TITLE OF INVENTION: For The Treatment Of Tumors
; FILE REFERENCE: A2778A-US
; CURRENT APPLICATION NUMBER: US/09/403,736
; CURRENT FILING DATE: 1999-10-26
; PRIOR APPLICATION NUMBER: PCT/EP98/02491
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/044,980
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: humanurokinase
;
; US-09-403-736-2
;
; Query Match 100.0%; Score 554; DB 4; Length 411;
; Best Local Similarity 100.0%; Pred. No. 7.2e-58;
; Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 KTCYEGNGHFGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 60
; DB 48 KTCYEGNGHFGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 107
;
; QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPSPPEE 96
; DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPSPPEE 143
;
; RESULT 12
; US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
```

CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/07/942.157A
FILING DATE: 19920908
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/631673
FILING DATE: 20-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: TS1108Cont.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)815-6508
TELEFAX: (404)815-6555
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..19
OTHER INFORMATION: /label= peptide
OTHER INFORMATION: /note= "WAP signal"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 198..203
OTHER INFORMATION: /label= modified
OTHER INFORMATION: /note= "six amino acids deleted in mutant"
US-07-942-157A-3
Query Match 100.0%; Score 554; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 7.6e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 67 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 126
Qy 61 RRRPWCYVQGLKPLVQECMVHDCADGKPSPPPEE 96
Db 127 RRRPWCYVQGLKPLVQECMVHDCADGKPSPPPEE 162
RESULT 13
US-09-101-272G-1
Sequence 1, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: Q50979
CURRENT APPLICATION NUMBER: US/09/101.272G
CURRENT FILING DATE: 1998-07-08
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patent in version 3.1
SEQ ID NO 1
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: mat_peptide
LOCATION: (21)..()
OTHER INFORMATION:
NAME/KEY: misc_feature
LOCATION: (20)..()
OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)
US-09-101-272G-1
Query Match 100.0%; Score 554; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.6e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 127
Qy 61 RRRPWCYVQGLKPLVQECMVHDCADGKPSPPPEE 96
Db 128 RRRPWCYVQGLKPLVQECMVHDCADGKPSPPPEE 163
RESULT 14
5188829-1
Patent No. 5188829
APPLICANT: KOBAYASHI, YO-ICHI-OMORI, MUNAKI, YAMADA, CHIKAKO
TITLE OF INVENTION: RAPIDLY ACTING PROTEINASE
NUMBER OF SEQUENCES: 23
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/340.007
FILING DATE: 18-AUG-1988
SEQ ID NO: 1
LENGTH: 431
5188829-1
Query Match 100.0%; Score 554; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.6e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 127
Qy 61 RRRPWCYVQGLKPLVQECMVHDCADGKPSPPPEE 96
Db 128 RRRPWCYVQGLKPLVQECMVHDCADGKPSPPPEE 163
RESULT 15
US-08-560-098A-47
Sequence 47, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENDEL, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560.098A

```
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-47:
;
; Query Match 100.0%; Score 554; DB 2; Length 432;
; Best Local Similarity 100.0%; Pred. No. 7.6e-58; Indels 0; Gaps 0;
; Matches 95; Conservative 0; Mismatches 0;
;
; QY 1 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQQTTHAHRSDALQLGLGKHYCRNPDN 60
; DB 69 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQQTTHAHRSDALQLGLGKHYCRNPDN 128
;
; QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
; DB 129 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 164
;
; RESULT 16
; US-09-181-816-1
; Sequence 1, Application US/09181816
; Patent No. 6277818
; GENERAL INFORMATION:
; APPLICANT: MAZAR, Andrew P.
; APPLICANT: JONES, Terence R.
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
; FILE REFERENCE: 329042000300 SIDN 1-7
; CURRENT APPLICATION NUMBER: US/09/181,816
; CURRENT FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-181-816-1
;
; Query Match 98.2%; Score 544; DB 3; Length 411;
; Best Local Similarity 99.0%; Pred. No. 1.1e-56; Indels 1; Gaps 0;
; Matches 95; Conservative 0; Mismatches 1;
;
; QY 1 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQQTTHAHRSDALQLGLGKHYCRNPDN 60
; DB 48 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQQTTHAHRSDALQLGLGKHYCRNPDN 107
;
; QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
; DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 143
;
; RESULT 17
; US-08-560-098A-48
; Sequence 48, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENDEL, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
```

```
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins Having Fibrinolytic and
; APPLICATION NUMBER: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-48
;
; Query Match 98.0%; Score 543; DB 2; Length 411;
; Best Local Similarity 99.0%; Pred. No. 1.5e-56; Indels 1; Gaps 0;
; Matches 95; Conservative 0; Mismatches 1;
;
; QY 1 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQQTTHAHRSDALQLGLGKHYCRNPDN 60
; DB 48 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQQTTHAHRSDALQLGLGKHYCRNPDN 107
;
; QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
; DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 143
;
; RESULT 18
; 5219569-2
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO: 2
; LENGTH: 430
; 5219569-2
;
; Query Match 97.2%; Score 538.5; DB 6; Length 430;
; Best Local Similarity 99.0%; Pred. No. 5.3e-56; Indels 1; Gaps 1;
; Matches 95; Conservative 0; Mismatches 0;
;
; QY 1 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQQTTHAHRSDALQLGLGKHYCRNPDN 60
```

Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNPDN 127
Qy 61 RRRPWCYVQVGLKPLVQECMWHDCADGKXPSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECMWHDCADGKXPSPPEE 162

RESULT 19
US-08-142-590B-25
; Sequence 25, Application US/08142590B
; Patent No. 6120765
; GENERAL INFORMATION:
; APPLICANT: HIBINO, Taahiko, TAKAHASHI, Tadahito; HORII, Izumi; and GORTINCK,
; TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,590B
; FILING DATE: 25-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,318
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: MGP-009CP
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-142-590B-25

Query Match 95.78; Score 530; DB 3; Length 157;
Best Local Similarity 95.8; Pred. No. 1.6e-55;
Matches 92; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNPDN 60
Db 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNPDN 107

Qy 61 RRRPWCYVQVGLKPLVQECMWHDCADGKXPSPPEE 96
Db 108 RRRPWCYVQVGLKPLVQECMWHDCADGKXPSPPEE 143

RESULT 20
US-08-797-689-12
; Sequence 12, Application US/08797689
; Patent No. 5878969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice

; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-797-689-12

Query Match 92.1; Score 510; DB 2; Length 138;
Best Local Similarity 100.0; Pred. No. 3.4e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNPDN 60
Db 51 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNPDN 110

Qy 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88
Db 111 RRRPWCYVQVGLKPLVQECMWHDCADGK 138

RESULT 21
US-09-984-186-12
; Sequence 12, Application US/09984186
; Patent No. 6686179
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43

CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12

Query Match 92.1%; Score 510; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.4e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 51 KTCYEGNGHYRGKASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 110
QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88
DB 111 RRRPWCYVQVGLKPLVQECWVHDCADGK 138

RESULT 22
US-09-101-272G-80
; Sequence 80, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 80
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: residues 43-131 of the ATF domain of uPA
US-09-101-272G-80

Query Match 88.3%; Score 489; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 6.3e-51;

Query Match 91.2%; Score 505; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 2e-52;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 49 KTCYEGNGHYRGKASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 108
QY 61 RRRPWCYVQVGLKPLVQECWVHDCADG 87
DB 109 RRRPWCYVQVGLKPLVQECWVHDCADG 135

RESULT 23
US-09-101-272G-96
; Sequence 96, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 96
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFHI-CL chimeric protein
US-09-101-272G-96

Query Match 91.2%; Score 505; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.1e-52;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 49 KTCYEGNGHYRGKASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 108
QY 61 RRRPWCYVQVGLKPLVQECWVHDCADG 87
DB 109 RRRPWCYVQVGLKPLVQECWVHDCADG 135

RESULT 24
US-09-101-272G-62
; Sequence 62, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 62
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: residues 43-131 of the ATF domain of uPA
US-09-101-272G-62

Query Match 88.3%; Score 489; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 6.3e-51;

```
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQLGLGKHNCRPN 60
Db 6 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQLGLGKHNCRPN 65
QY 61 RRPWCYVQVGLKPLVQECMVHDC 84
Db 66 RRPWCYVQVGLKPLVQECMVHDC 89

RESULT 25
US-08-560-098A-51
; Sequence 51, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEPFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-51

Query Match 43.5%; Score 241; DB 2; Length 477;
Best Local Similarity 50.0%; Pred. No. 1.7e-20;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;
QY 2 TCYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQLGLGKHNCRPN 61
Db 127 TCYEQGVYRGVTSSTESGACINWNSNLLTRTYNGRRSDAITLGLGHNCRPN 186
QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
Db 187 SKPWCYVFKASKFILEFCSPVCS 210

RESULT 26
US-07-609-510B-16
; Sequence 16, Application US/07609510B
; Patent No. 5326700
```

```
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue PI
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,510B
; FILING DATE: 19901106
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 40.8%; Score 226; DB 1; Length 527;
Best Local Similarity 47.7%; Pred. No. 1.1e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
QY 2 TCYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQLGLGKHNCRPN 61
Db 91 TCYEDQGISYRGVTSSTESGACINWNSNLLTRTYNGRRSDAITLGLGHNCRPN 150
QY 62 RRPWCYVQVGLKPLVQECMVHDC 87
Db 151 SKPWCYVFKACKYSSEFCSTPACSEG 176

RESULT 27
US-08-811-949-39
; Sequence 39, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```


Db 91 TCYEDQGISYRGWTSAESGAECTNNSSALAQKPSGRRPDAIRLGLGNHNYCRNPDNR 150
QY 62 RRPWCYVQVGLKPLVQECVHDCADG 87
Db 151 SKPCWCVFKAGKYSSEFCSTPACSEG 176

RESULT 31
US-08-811-949-43
Patent No. 5200340-6
APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,
PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINGJI
TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
ACTIVATORS
NUMBER OF SEQUENCES: 34
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/53,412
FILING DATE: 22-MAY-1987
SEQ ID NO:6:
LENGTH: 546
5200340-6

Query Match 40.8%; Score 226; DB 6; Length 546;
Best Local Similarity 47.7%; Pred. No. 1.2e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRNPDNR 61
Db 126 TCYEDQGISYRGWTSAESGAECTNNSSALAQKPSGRRPDAIRLGLGNHNYCRNPDNR 185
QY 62 RRPWCYVQVGLKPLVQECVHDCADG 87
Db 186 SKPCWCVFKAGKYSSEFCSTPACSEG 211

RESULT 32
US-08-811-949-43
Sequence 43, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINRO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-43
Query Match 40.8%; Score 226; DB 2; Length 562;
Best Local Similarity 47.7%; Pred. No. 1.2e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRNPDNR 61
Db 126 TCYEDQGISYRGWTSAESGAECTNNSSALAQKPSGRRPDAIRLGLGNHNYCRNPDNR 185
QY 62 RRPWCYVQVGLKPLVQECVHDCADG 87
Db 186 SKPCWCVFKAGKYSSEFCSTPACSEG 211

RESULT 33
US-08-560-098A-50
Sequence 50, Application US/08560098A
Patent No. 5576841
GENERAL INFORMATION:
APPLICANT: WENDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-50

Query Match 40.8%; Score 226; DB 2; Length 562;
Best Local Similarity 47.7%; Pred. No. 1.2e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRNPDNR 61
Db 126 TCYEDQGISYRGWTSAESGAECTNNSSALAQKPSGRRPDAIRLGLGNHNYCRNPDNR 185

```
QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
:|||||:
DB 186 SKPCYVFKAGKYSSEFCSTPACSEG 211

RESULT 34
US-08-883-795A-38
; Sequence 38, Application US/08883795A
; Patent No. 5985607
; GENERAL INFORMATION:
; APPLICANT: Delcove, Genevieve
; APPLICANT: Awang, Gregor
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERSKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,795A
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 7841-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 562 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein (tpa)
US-08-883-795A-38

Query Match 40.8%; Score 226; DB 2; Length 562;
Best Local Similarity 47.7%; Pred. No. 1.2e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDNR 61
:|||||:
DB 126 TCYEDQGISYRGTWSTAESGAECTNNSSALAKPYSGRRPDAIRLGLGNHNYCRNPDNR 185

QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
:|||||:
DB 186 SKPCYVFKAGKYSSEFCSTPACSEG 211

RESULT 35
US-09-703-695A-4
; Sequence 4, Application US/09703695A
; Patent No. 6593097
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: P1788R1
; CURRENT APPLICATION NUMBER: US/09/703,695A
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04

Query Match 40.8%; Score 226; DB 6; Length 562;
Best Local Similarity 47.7%; Pred. No. 1.2e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDNR 61
:|||||:
DB 126 TCYEDQGISYRGTWSTAESGAECTNNSSALAKPYSGRRPDAIRLGLGNHNYCRNPDNR 185

QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
:|||||:
DB 186 SKPCYVFKAGKYSSEFCSTPACSEG 211

RESULT 36
US-08-883-795A-38
; Patent No. 5185259
; APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHAR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO: 3:
; LENGTH: 562
5185259-3

Query Match 40.8%; Score 226; DB 6; Length 562;
Best Local Similarity 47.7%; Pred. No. 1.2e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDNR 61
:|||||:
DB 126 TCYEDQGISYRGTWSTAESGAECTNNSSALAKPYSGRRPDAIRLGLGNHNYCRNPDNR 185

QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
:|||||:
DB 186 SKPCYVFKAGKYSSEFCSTPACSEG 211

RESULT 37
US-09-703-695A-4
; Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,
; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
; ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
; SEQ ID NO: 2:
; LENGTH: 562
5200340-2

Query Match 40.8%; Score 226; DB 6; Length 562;
```

Best Local Similarity 47.7%; Pred. No. 1.2e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDR 61
DB 126 TCYEDQGISYRGTWSTAESGAECTWNSSALQPKYSGRRPDAIRLGLGNHNYCRNPDRD 185

QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
DB 186 SKPCYVFKAGKYSSEFCSTPACSEG 211

RESULT 38
5344773-2
; Patent No. 5344773
; APPLICANT: WEI, CHA-MER; HSIUNG, NANCY; REDDY, VERMURI B.;
; LEWONTT, JEFFREY P.; DACKOWSKI, WILLIAM; DOUGLAS, RICHARD;
; COLE, EDWARD S.; PURCELL JR., RICHARD D.; LAU, DAVID TAI-YUI
; TITLE OF INVENTION: HUMAN UTERINE TISSUE PLASMINOGEN
; ACTIVATOR PRODUCED BY RECOMBIANT DNA
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/782,686
; FILING DATE: 01-OCT-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 656,770
; FILING DATE: 01-OCT-1984
; SEQ ID NO: 2:
; LENGTH: 562
5344773-2

Query Match 40.8%; Score 226; DB 6; Length 562;
Best Local Similarity 47.7%; Pred. No. 1.2e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDR 61
DB 126 TCYEDQGISYRGTWSTAESGAECTWNSSALQPKYSGRRPDAIRLGLGNHNYCRNPDRD 185

QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
DB 186 SKPCYVFKAGKYSSEFCSTPACSEG 211

RESULT 39
US-08-427-640-8
; Sequence 8, Application US/08427640
; Patent No. 5658788
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Compounds and Methods for Treatment of Thromboembolic Disor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN.
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,640
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/689,410
; FILING DATE: 22 APRIL 1991
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 356 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-427-640-8

Query Match 40.0%; Score 221.5; DB 1; Length 356;
Best Local Similarity 44.2%; Pred. No. 2.4e-18;
Matches 42; Conservative 12; Mismatches 40; Indels 1; Gaps 1;

QY 2 TCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDR 61
DB 8 TCYEDQGISYRGTWSTAESGAECTWNSSALQPKYSGRRPDAIRLGLGNHNYCRNPDRD 67

QY 62 RRPWCYVQVGLKPLVQECMVHDCAD-GKKPSSPPE 95
DB 68 AKPWCHLVKNNRLTWEYCDVPSCSTCGLRQYSQPQ 102

RESULT 40
US-08-811-949-49
; Sequence 49, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-49

Query Match 39.9%; Score 221; DB 2; Length 437;
Best Local Similarity 47.1%; Pred. No. 3.6e-18;
Matches 40; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDR 62
DB 2 CYEDQGISYRGTWSTAESGAECTWNSSALQPKYSGRRPDAIRLGLGNHNYCRNPDRD 61

QY 63 RRPWCYVQVGLKPLVQECMVHDCADG 87

Db 62 KPWCYVFKAGKYSSEFCSTPACSEG 86

RESULT 41

US-08-811-949-51
; Sequence 51, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-51

Query Match 39.9%; Score 221; DB 2; Length 437;
Best Local Similarity 47.1%; Pred. No. 3.6e-18;
Matches 40; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
QY 3 CYEGNGHFYRGKASTDTMGPRCLPWSATVLQOTYHAHRSALQLGLGKHNYCNPDR 62
Db 2 CYEDQGISYRGTWSTAESGAECTNWSALAQAQPSYSGRRPDATRLGLGNHNYCNPDRDS 61
QY 63 RPWCYVQVGLKPLVQECMVHDCADG 87
Db 62 KPWCYVFKAGKYSSEFCSTPACSEG 86

RESULT 42

US-08-811-949-55
; Sequence 55, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU

; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-55

Query Match 39.9%; Score 221; DB 2; Length 437;
Best Local Similarity 47.1%; Pred. No. 3.6e-18;
Matches 40; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
QY 3 CYEGNGHFYRGKASTDTMGPRCLPWSATVLQOTYHAHRSALQLGLGKHNYCNPDR 62
Db 2 CYEDQGISYRGTWSTAESGAECTNWSALAQAQPSYSGRRPDATRLGLGNHNYCNPDRDS 61
QY 63 RPWCYVQVGLKPLVQECMVHDCADG 87
Db 62 KPWCYVFKAGKYSSEFCSTPACSEG 86

RESULT 43

US-08-811-949-57
; Sequence 57, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/811.949
;; FILING DATE: 05-MAR-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: OBLON, NORMAN F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 18-966-0
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-413-3000
;; TELEFAX: 703-413-2220
;; INFORMATION FOR SEQ ID NO: 57:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 437 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: Protein
US-08-811-949-57

Query Match 39.9%; Score 221; DB 2; Length 437;
Best Local Similarity 47.1%; Pred. No. 3.5e-18;
Matches 40; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
QY 3 CYEGNGHYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHYCRNPDNR 62
Db 2 CYEDGHSYRGTSHTAESGACLPWNSALAKPYSGRRPDAIRLGLGNHNYCRNPDNR 61
QY 63 RPMCYYVQGLKPLVQECMVHDCADG 87
Db 62 KPCVYVFKAGKYSSEFCSTPACSEG 86

RESULT 44
US-09-553-498-10
; Sequence 10, Application US/09553498
; Patent No. 6309861
; GENERAL INFORMATION:
; APPLICANT: Ambrosius, Dorothee
; APPLICANT: Rudolph, Rainer
; APPLICANT: Schaeffner, Joerg
; APPLICANT: Schwarz, Elisabeth
; TITLE OF INVENTION: Process for the production of naturally folded and secreted prote
; FILE REFERENCE: Case 20379
; CURRENT APPLICATION NUMBER: US/09/553.498
; CURRENT FILING DATE: 2000-04-20
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 10
; LENGTH: 378
; TYPE: PRT
; ORGANISM: E. coli
US-09-553-498-10

Query Match 39.8%; Score 220.5; DB 4; Length 378;
Best Local Similarity 44.7%; Pred. No. 3.5e-18;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;
QY 3 CYEGNGHYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHYCRNPDNR 62
Db 31 CYFNGSAYRGTHSLTESGASCLPWSMILIGKYVTAQNPSAQLGLGKHYCRNPDGDA 90
QY 63 RPMCYYVQGLKPLVQECMVHDCADG-GKXPSPPE 95
Db 91 KPWCVLTNRRLTWECYDVPSCSTCGLRQYSQPQ 124

RESULT 45
US-09-618-869-10
; Sequence 10, Application US/09618869
; Patent No. 645279
; GENERAL INFORMATION:
; APPLICANT: Ambrosius, Dorothee

;; APPLICANT: Rudolph, Rainer
;; APPLICANT: Schaeffner, Joerg
;; APPLICANT: Schwarz, Elisabeth
;; TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND
;; TITLE OF INVENTION: SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR
;; TITLE OF INVENTION: CHAPERONES
;; FILE REFERENCE: 20381
;; CURRENT APPLICATION NUMBER: US/09/618.869
;; CURRENT FILING DATE: 2000-07-19
;; PRIOR APPLICATION NUMBER: EP99114811.5
;; PRIOR FILING DATE: 1999-07-29
;; NUMBER OF SEQ ID NOS: 10
;; SOFTWARE: Patent in Ver. 2.1
;; SEQ ID NO 10
;; LENGTH: 378
;; TYPE: PRT
;; ORGANISM: Escherichia coli
US-09-618-869-10

Query Match 39.8%; Score 220.5; DB 4; Length 378;
Best Local Similarity 44.7%; Pred. No. 3.5e-18;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;
QY 3 CYEGNGHYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHYCRNPDNR 62
Db 31 CYFNGSAYRGTHSLTESGASCLPWSMILIGKYVTAQNPSAQLGLGKHYCRNPDGDA 90
QY 63 RPMCYYVQGLKPLVQECMVHDCADG-GKXPSPPE 95
Db 91 KPWCVLTNRRLTWECYDVPSCSTCGLRQYSQPQ 124

RESULT 46
US-08-811-949-67
; Sequence 67, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C. SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811.949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid

TOPOLGY: linear
MOLECULE TYPE: protein
US-08-811-949-67

Query Match 39.6%; Score 219.5; DB 2; Length 389;
Best Local Similarity 44.7%; Pred. No. 4.7e-18;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;
Qy 3 CYEGNGHFRGKASTDTMGRCPLPWNATVLOQTYHAHRSDALQGLGKHNCRPNRR 62
Db 42 CYFGNSAYRGTHSLTESGASCLPWNAMILIGKVTYQNPQAALGLGKHNCRPDGDA 101
Qy 63 RWCYVQVGLKPLVQECWHDCA-DGKFPSSPPE 95
Db 102 KPWCHLVKRRLTWEYCDVPSCTGLRQYSQPQ 135

RESULT 47

US-09-411-977-3

Sequence 3, Application US/09411977
Patent No. 6372473
GENERAL INFORMATION:
APPLICANT: Moore, Paul A.
APPLICANT: Ruben, Steven M.
APPLICANT: Eber, Reinhard
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
FILE REFERENCE: PF378P1
CURRENT APPLICATION NUMBER: US/09/411,977
CURRENT FILING DATE: 1999-10-04
EARLIER APPLICATION NUMBER: 09/084,491
EARLIER FILING DATE: 1998-05-27
EARLIER APPLICATION NUMBER: 60/048,000
EARLIER FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 326
TYPE: PRT
ORGANISM: Homo sapiens
US-09-411-977-3

Query Match 39.4%; Score 218.5; DB 4; Length 326;
Best Local Similarity 44.7%; Pred. No. 5e-18;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;
Qy 3 CYEGNGHFRGKASTDTMGRCPLPWNATVLOQTYHAHRSDALQGLGKHNCRPNRR 62
Db 25 CYFGNSAYRGTHSLTESGASCLPWNAMILIGKVTYQNPQAALGLGKHNCRPDGDA 84
Qy 63 RWCYVQVGLKPLVQECWHDCA-DGKFPSSPPE 95
Db 85 KPWCHLVKRRLTWEYCDVPSCTGLRQYSQPQ 118

RESULT 48

US-08-137-116-1

Sequence 1, Application US/08137116
Patent No. 5500411
GENERAL INFORMATION:
APPLICANT: Martin, Ulrich
APPLICANT: Koshig, Reinhard
TITLE OF INVENTION: METHOD FOR TREATING THROMBOEMBOLIC
TITLE OF INVENTION: CONDITIONS BY INHIBITING REOCCUSION
TITLE OF INVENTION: VIA THE USE OF MULTIPLE BOLUS
TITLE OF INVENTION: ADMINISTRATION OF THROMBOLYTICALLY
TITLE OF INVENTION: ACTIVE PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA

ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,116
FILING DATE: June 30, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP92/00851
FILING DATE: 15 April 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 41 12 398
FILING DATE: 16 April 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 41 23 845
FILING DATE: 18 July 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5500411man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: ROER 1026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acids
TOPOLOGY: linear
US-08-137-116-1

Query Match 39.4%; Score 218.5; DB 1; Length 355;
Best Local Similarity 44.7%; Pred. No. 5.6e-18;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;
Qy 3 CYEGNGHFRGKASTDTMGRCPLPWNATVLOQTYHAHRSDALQGLGKHNCRPNRR 62
Db 8 CYFGNSAYRGTHSLTESGASCLPWNAMILIGKVTYQNPQAALGLGKHNCRPDGDA 67
Qy 63 RWCYVQVGLKPLVQECWHDCA-DGKFPSSPPE 95
Db 68 KPWCHLVKRRLTWEYCDVPSCTGLRQYSQPQ 101

RESULT 49

US-08-217-618-1

Sequence 1, Application US/08217618
Patent No. 5510330
GENERAL INFORMATION:
APPLICANT: Martin, Ulrich
APPLICANT: Fischer, Stephan
TITLE OF INVENTION: COMBINATIONS OF THROMBOLYTICALLY ACTIVE
TITLE OF INVENTION: PROTEINS AND NON-HEPARIN ANTICOAGULANTS, AND USES THEREOF
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/217,618
FILING DATE: 25-MARCH-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:

Search completed: May 25, 2004, 15:00:09
Job time : 7.6468 secs

NAME: Hanson, No. 5510330man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: BOER 1032
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acids
TOPOLOGY: linear
US-08-217-618-1

Query Match 39.4%; Score 218.5; DB 1; Length 355;
Best Local Similarity 44.7%; Pred. No. 5.6e-18;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;
Qy 3 CYEGNGHYFGKASTDTMGRCPLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDR 62
Db 8 CYFGNGSAYRTHSLTESGASCLPWSMILIGKYTAQNPSAQLGLGKHNYCRNP 67
Qy 63 RPYCVQVGLKPLVOECMWHDCAD-GKKPSPP 95
Db 68 KPCHVLKRRRLTWECYDVPSCSTCGLRQYSQP 101

RESULT 50
US-08-427-640-2
; Sequence 2, Application US/08427640
; Patent No. 5658788
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Compounds and Methods for Treatment of Thromboembolic Disor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/427,640
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/689,410
FILING DATE: 22 APRIL 1991
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-427-640-2

Query Match 39.4%; Score 218.5; DB 1; Length 355;
Best Local Similarity 44.7%; Pred. No. 5.6e-18;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;
Qy 3 CYEGNGHYFGKASTDTMGRCPLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDR 62
Db 8 CYFGNGSAYRTHSLTESGASCLPWSMILIGKYTAQNPSAQLGLGKHNYCRNP 67
Qy 63 RPYCVQVGLKPLVOECMWHDCAD-GKKPSPP 95
Db 68 KPCHVLKRRRLTWECYDVPSCSTCGLRQYSQP 101